Single-cell Raman spectroscopy of irradiated tumour cells

by

Quinn Matthews
B.Sc., University of Victoria, 2006
M.Sc., University of Victoria, 2008

A Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of

DOCTOR OF PHILOSOPHY

in the Department of Physics and Astronomy

Single-cell Raman spectroscopy of irradiated tumour cells

by

Quinn Matthews
B.Sc., University of Victoria, 2006
M.Sc., University of Victoria, 2008

Supervisory Committee

Dr. A. Jirasek, Supervisor
(Department of Physics and Astronomy)

Dr. M. Lefebvre, Member
(Department of Physics and Astronomy)

Dr. W. Ansbacher, Member
(Department of Physics and Astronomy; British Columbia Cancer Agency - Vancouver Island Centre)

Dr. A. G. Brolo, Outside Member
(Department of Chemistry - University of Victoria)

Supervisory Committee

Dr. A. Jirasek, Supervisor
(Department of Physics and Astronomy)

Dr. M. Lefebvre, Member
(Department of Physics and Astronomy)

Dr. W. Ansbacher, Member
(Department of Physics and Astronomy; British Columbia Cancer Agency - Vancouver Island Centre)

Dr. A. G. Brolo, Outside Member
(Department of Chemistry - University of Victoria)

ABSTRACT

This work describes the development and application of a novel combination of single-cell Raman spectroscopy (RS), automated data processing, and principal component analysis (PCA) for investigating radiation induced biochemical responses in human tumour cells. The developed techniques are first validated for the analysis of large data sets (∼200 spectra) obtained from single cells. The effectiveness and robustness of the automated data processing methods is demonstrated, and potential pitfalls that may arise during the implementation of such methods are identified. The techniques are first applied to investigate the inherent sources of spectral variability between single cells of a human prostate tumour cell line (DU145) cultured *in vitro*. PCA is used to identify spectral differences that correlate with cell cycle progression and the changing confluency of a cell culture during the first 3-4 days after sub-culturing. Spectral variability arising from cell cycle progression is (i) expressed as varying intensities of protein and nucleic acid features relative to lipid features,

(ii) well correlated with known biochemical changes in cells as they progress through the cell cycle, and (iii) shown to be the most significant source of inherent spectral variability between cells. This characterization provides a foundation for interpreting spectral variability in subsequent studies. The techniques are then applied to study the effects of ionizing radiation on human tumour cells. DU145 cells are cultured *in vitro* and irradiated to doses between 15 and 50 Gy with single fractions of 6 MV photons from a medical linear accelerator. Raman spectra are acquired from irradiated and unirradiated cells, up to 5 days post-irradiation. PCA is used to distinguish radiation induced spectral changes from inherent sources of spectral variability, such as those arising from cell cycle. Radiation induced spectral changes are found to correlate with both the irradiated dose and the incubation time post-irradiation, and to arise from biochemical differences in lipids, nucleic acids, amino acids, and conformational protein structures between irradiated and unirradiated cells. This study is the first use of RS to observe radiation induced biochemical effects in single cells, and is the first use of vibrational spectroscopy to observe such effects independent from cell cycle or cell death related processes. The same methods are then applied to a panel of human tumour cell lines, derived from prostate (DU145, PC3, LNCaP and PacMet), breast (MDA-MB-231 and MCF7) and lung (H460), which vary by p53 gene status and intrinsic radiosensitivity. One radiation induced PCA component is detected for each cell line by statistically significant changes in the PCA score distributions for irradiated samples, as compared to unirradiated samples, in the first 24 to 72 hours post-irradiation. These RS response signatures arise from radiation induced changes in cellular concentrations of aromatic amino acids, conformational protein structures, and certain nucleic acid and lipid functional groups. Correlation analysis between the radiation induced PCA components separates the cell lines into three unique RS response categories: R1 (H460, MCF7 and PacMet), R2 (MDA-MB-231 and PC3), and R3 (DU145 and LNCaP). These RS categories partially segregate according to radiosensitivity; the R1 and R2 cell lines are radioresistant and the R3 cell lines are radiosensitive (PacMet radiosensitivity (R1) unknown). The R1 and R2 cell lines further segregate according to p53 gene status, corroborated by cell cycle analysis post-irradiation. Preliminary results obtained from a mouse prostate tumour cell line (TRAMP-C2), irradiated both *in vitro* and *in vivo*, indicate that RS signatures of radiation response may also be detectable from tumour cells obtained from an *in vivo* system during radiation therapy treatment. These results indicate the potential for future RS studies designed to investigate, monitor, or predict radiation response.

Contents

Supervisory Committee ii

Abstract iii

Table of Contents vi

List of Tables xi

List of Figures xiii

Acknowledgements xviii

Acknowledgements xx

1 Introduction 1

1.1 Radiation therapy . . . 1

1.1.1 History . . . 1

1.1.2 Modern clinical implementation . . . 2

1.1.3 External beam radiation therapy . . . 3

1.2 Cellular radiobiology . . . 7

1.2.1 The human cell: Biomolecules, cell cycle and genes . . . 7

1.2.2 Radiation effects on biomolecules . . . 11

1.2.3 Radiation effects on cell survival . . . 13

1.2.4 Current problems and questions in radiobiology . . . 16

1.2.5 Previous and current efforts for predicting or monitoring tumour response and intrinsic radiosensitivity . . . 18

1.3 Raman spectroscopy in cell and tissue analysis . . . 21

1.3.1 Advantages of Raman spectroscopy for biomedical applications 22

1.3.2 Raman spectroscopy of cells and tissues . . . 22

1.3.3 Raman spectroscopy in radiobiology . . . 23

1.4 Thesis scope 24

2 Raman Spectroscopy **27**

2.1 History 27
2.2 Theory of Raman scattering 28
2.2.1 Origin of the Raman effect 28
2.2.2 Intensity of Raman scattering 32
2.2.3 Molecular vibrations and Raman activity 34
2.3 Raman spectroscopy instrumentation 37
2.3.1 Raman shift 37
2.3.2 Raman spectroscopy apparatus 37
2.3.3 Light dispersion 38
2.3.4 Light detection and Raman spectrum creation 40
2.3.5 Raman microscopy 42
2.4 Raman spectroscopy of cells 45
2.4.1 Laser wavelength and laser power 45
2.4.2 Spatial and confocal resolution 47
2.4.3 Spectral windows and spectral resolution 47
2.4.4 Substrate material 47
2.5 Advantages and disadvantages of Raman spectroscopy 48
2.5.1 General advantages 48
2.5.2 General disadvantages 48
2.5.3 General considerations for biological samples 49
2.5.4 Comparisons with established molecular analysis techniques for cells or tissues 50
2.6 Example applications of Raman spectroscopy 52

3 Materials & Methods **54**

3.1 Cells and cell processing 54
3.1.1 Human tumour cells 54
3.1.2 Cell culture 55
3.1.3 Sample preparation for Raman spectroscopy 57
3.1.4 Cell irradiation experiments 57
3.1.5 Flow cytometry analysis of cell cycle and viability 59
3.2 Raman spectroscopy 60

3.2.1 InVia Raman microscope 60
3.2.2 Single-cell spectral acquisition 63
3.3 Spectral processing 64
3.3.1 Cosmic ray removal 64
3.3.2 Two-point maximum entropy method smoothing 65
3.3.3 Baseline estimation 66
3.3.4 Spectral normalization 70
3.3.5 Principal component analysis 72

4 Results & Discussion I: Validation of Analysis Methods 77
4.1 Introduction 77
4.2 Spectral smoothing 78
4.2.1 Effect on processed spectra 78
4.2.2 Effect on PCA components and PCA scores 79
4.3 Baseline estimation 83
4.3.1 Effect on processed spectra 83
4.3.2 Effect on PCA components and PCA scores 85
4.4 Spectral normalization 91
4.4.1 Effect on processed spectra 92
4.4.2 Effect on PCA components and PCA scores 94
4.5 Discussion 99
4.6 Conclusion 101

5 Results & Discussion II: Variability in Raman spectra of single human tumour cells 102
5.1 Introduction 102
5.2 Materials & Methods 104
5.2.1 Cell preparation 104
5.2.2 Raman spectroscopy and data processing 105
5.3 Results 105
5.3.1 Single DU145 cell spectrum 105
5.3.2 Study #1: Asynchronous cell cultures 108
5.3.3 Study #2: Synchronized cell cultures 117
5.4 Discussion 123

5.4.1 Study #1: Asynchronous cell cultures . . . 123
5.4.2 Study #2: Synchronized cell cultures . . . 126
5.4.3 Spectral variability and PCA . . . 129
5.4.4 LWN vs. HWN spectral windows . . . 131
5.4.5 Spectral variability and cell size . . . 131
5.5 Conclusion . . . 133

6 Results & Discussion III: Raman spectroscopy of single human tumour cells exposed to ionizing radiation 134
6.1 Introduction . . . 134
6.2 Methods . . . 135
6.3 Results . . . 136
6.3.1 Single DU145 cell spectrum . . . 136
6.3.2 Irradiated vs. unirradiated cells . . . 137
6.3.3 Effect of time of irradiation after sub-culturing . . . 145
6.4 Discussion . . . 147
6.4.1 First PCA component: Cell cycle variability . . . 147
6.4.2 Second PCA component: Radiation induced effects . . . 150
6.4.3 LWN vs. HWN spectral windows . . . 153
6.5 Conclusion . . . 154

7 Results & Discussion IV: Biochemical signatures of radiation response in lung, breast and prostate tumour cells 155
7.1 Introduction . . . 155
7.2 Materials & Methods: Study #1 . . . 157
7.3 Results: Study #1 . . . 158
7.3.1 Unirradiated cell spectra . . . 158
7.3.2 Cell cycle spectral variability . . . 158
7.3.3 RS radiation response signatures I: Categories R1, R2 and R3 160
7.3.4 RS radiation response signatures II: Radiation induced changes in biomolecules across categories R1, R2 and R3 . . . 161
7.3.5 RS radiation response associations with cell cycle arrest, p53 and radiosensitivity . . . 165

7.4 Discussion: Study #1 165
7.4.1 RS detection of biochemical signatures of radiation response . 165
7.4.2 Segregation of common radiation response signatures according to p53 status and radiosensitivity 167
7.4.3 Biochemical mechanisms of radiation resistance or sensitivity . 168
7.4.4 Uniqueness of the observed RS biochemical radiation responses 170
7.4.5 Effect of radiation on cell cycle variability 171
7.5 Materials & Methods: Study #2 171
7.5.1 TRAMP-C2 cells 171
7.5.2 *In vitro* experiment 172
7.5.3 *In vivo* experiment 172
7.6 Results & Discussion: Study #2 173
7.6.1 *In vitro* experiment 173
7.6.2 *In vivo* experiment 179
7.6.3 Combined *in vitro* and *in vivo* data set analysis 184
7.6.4 *In vivo* experimental difficulties and possible resolutions 188
7.7 Conclusion 189

8 Conclusions and Future Work **191**

Bibliography **196**

List of Tables

Table 1.1 Relative number of interactions (%) for photoelectric (τ), Compton (σ) and pair and triplet production (π) processes in water . 5

Table 2.1 Comparison of RS with flow cytometry, PET and MRS, for the molecular analysis of cells or tissues. 52

Table 3.1 Doubling time for the seven human tumour cell lines used in this work 55

Table 4.1 Absolute values of the relative distances between the mean values of adjacent PCA score distributions for the first two PCA components of the sample DU145 data set, using the SRM for BL estimation with single window sizes of 3%, 5% and 7% 90

Table 4.2 Absolute values of the relative distances between the mean values of adjacent PCA score distributions for the first two PCA components of the sample DU145 data set, when normalizing by the total area or by the area of the 1450 cm^{-1} peak 98

Table 5.1 Molecular assignments for spectra of DU145 cells 107

Table 6.1 Raman spectral changes correlated with dose (0 to 50 Gy) and post-irradiation incubation time (0 to 120 hours), observed in irradiated DU145 cells 151

Table 7.1 Correlation r-values between radiation induced PCA components for each human tumour cell line 161

Table 7.2 RS biochemical radiation response category, tissue of origin, percent variance explained by radiation induced PCA component, G1, S and G2 fractions at 24 hours post-irradiation, p53 status and average reported radiosensitivity for the seven cell lines used in this study 166

Table 7.3 Comparison of TRAMP-C2 *in vitro* RS biochemical radiation response category, tissue of origin, percent variance explained by radiation induced PCA component, G1, S and G2 fractions at 24 hours post-irradiation, p53 status and average reported radiosensitivity with the seven human tumour cell lines used in study #1 180

List of Figures

Figure 1.1 Schematic of a modern clinical linear accelerator for external beam radiation therapy 4
Figure 1.2 Schematic diagrams of (a) the general structure of an amino acid, (b) the joining of two amino acids in a dehydration reaction, and the secondary structure of crambin. 8
Figure 1.3 Block diagrams of double-stranded DNA and single-stranded RNA 9
Figure 1.4 Schematic diagram of phosphatidylcholine 10
Figure 1.5 Phases of the mitotic cell cycle 11
Figure 1.6 (a) Single dose survival curves and (b) fractionated dose survival curves 14
Figure 1.7 (a) General schematic of the light path for a typical RS system, and (b) a sample Raman spectrum 21

Figure 2.1 Energy level diagram for a molecule irradiated with optical photons of frequency ν_0 29
Figure 2.2 (a)-(c) Normal mode vibrations of the CO_2 molecule and (d) symmetric ring breathing of benzene 35
Figure 2.3 Changes in polarizability ellipsoids during normal mode vibrations of CO_2 36
Figure 2.4 Design of the dispersive spectrometer used in this work 39
Figure 2.5 The arrangement of a CCD detector array placed at the focal plane of the spectrometer 41
Figure 2.6 General schematic of the light path for a typical RM system . . 43
Figure 2.7 Principle of confocal Raman microscopy 44
Figure 2.8 Example of numerical apertures (NA) calculated for two microscope objectives 44

Figure 3.1 Optical images of human tumour cell lines 56

Figure 3.2 Protocol for irradiated dose and incubation time before RS analysis, for cell culture irradiation experiments 58
Figure 3.3 Linac set-up for cell culture irradiations 59
Figure 3.4 Light path through the inVia Raman microscopy system used in this work 61
Figure 3.5 Optical image of a portion of a cell pellet 63
Figure 3.6 Example of applying the TPMEM for smoothing of a raw Raman spectrum of a cell 66
Figure 3.7 Example of applying the SRM for BL estimation, showing the SG filter BL estimate and the SRM modified data for various iterations 68
Figure 3.8 Example of applying the SRM for baseline (BL) estimation, showing the effect of varying the SG filter window size (W) 69
Figure 3.9 Example of applying the SRM for BL estimation for a LWN spectral window cell spectrum, showing the method of using two different SG window sizes 70
Figure 3.10Example of applying the TPLI method for BL estimation for a HWN spectral window cell spectrum 71
Figure 3.11Example of applying normalization to four BL subtracted cell spectra 72
Figure 3.12Examples of PCA components from a 240 spectra LWN window data set used in this work 73
Figure 3.13Schematic of the PCA score plots for the first and second PCA components from a fictitious experiment 75

Figure 4.1 Effect of varying the TPMEM smoothing parameter X on two processed DU145 cell spectra 79
Figure 4.2 Effect of varying the TPMEM smoothing parameter X on the first two PCA components of the sample DU145 data set 81
Figure 4.3 Effect of varying the TPMEM smoothing parameter X on the PCA scores for the first two PCA components of the sample DU145 data set 82
Figure 4.4 Effect of using the SRM with single window sizes of (a) 3%, (b) 5% and (c) 7% on two processed cell spectra, and (d) the BL subtracted spectrum for a single cell using these window sizes . 84

Figure 4.5 Effect of using the SRM with single window sizes of 3%, 5% and 7% on the first two PCA components of the sample DU145 data set 86
Figure 4.6 Effect of using the SRM with single window sizes of 3%, 5% and 7% on the PCA scores for the first two PCA components of the sample DU145 data set 89
Figure 4.7 Effect of normalizing to the total area or the area under the 1450 cm^{-1} peak on two processed DU145 cell spectra 93
Figure 4.8 Effect of normalizing by the total area or by the area of the 1450 cm^{-1} peak on the first two PCA components of the sample DU145 data set 95
Figure 4.9 Effect of normalizing by the total area or by the area of the 1450 cm^{-1} peak on the PCA scores for the first two PCA components of the sample DU145 data set 97

Figure 5.1 Raman spectra of a single DU145 cell for the (a) LWN and (b) HWN spectral windows 106
Figure 5.2 Flow cytometry analysis of cell cycle distributions for the asynchronous cell cultures 108
Figure 5.3 First PCA components from the asynchronous cell cultures study 109
Figure 5.4 PCA scores for the first components from the asynchronous cell cultures study 111
Figure 5.5 Raman and difference spectra for two cells with a large difference in PCA score for the first PCA component 112
Figure 5.6 Second PCA components from the asynchronous cell cultures study 114
Figure 5.7 PCA scores for the second components from the asynchronous cell cultures study 114
Figure 5.8 Raman and difference spectra for two cells that have a large difference in PCA score for the second PCA component 115
Figure 5.9 Third PCA component and PCA scores, for the LWN spectral window, from the asynchronous cell cultures study 116
Figure 5.10 Flow cytometry analysis of cell cycle distributions for the synchronized cell cultures 118
Figure 5.11 First PCA components from the synchronized cell cultures study 119

Figure 5.12PCA scores for the first components from the synchronized cell cultures study 120
Figure 5.13Second PCA components from the synchronized cell cultures study 121
Figure 5.14PCA scores for the second components from the synchronized cell cultures study 122

Figure 6.1 Raman spectra and peak assignments for a single unirradiated DU145 cell 136
Figure 6.2 First PCA component results for the LWN window 139
Figure 6.3 Second PCA component results for the LWN window 141
Figure 6.4 Third PCA component and PCA scores for the LWN spectral window 142
Figure 6.5 First and second PCA component results for the HWN window 143
Figure 6.6 Second and third PCA component results for the LWN window, from a radiation experiment with cells irradiated at 20-30% confluency 146
Figure 6.7 Flow cytometry analysis of unirradiated (0 Gy, top row) and irradiated (50 Gy, bottom row) cell cycle distributions 148

Figure 7.1 (a) Sample Raman spectrum of a single unirradiated DU145 cell. (b) Average spectra from 20 unirradiated cells for the seven cell lines used in this study. (c) Cell cycle PCA component for the DU145 data set (200 cells). (d) Cell cycle PCA components for all seven cell lines 159
Figure 7.2 Radiation induced PCA components for all 7 cell lines 160
Figure 7.3 Radiation induced PCA components and PCA scores for the H460, MDA-MB-231 and DU145 cell lines 163
Figure 7.4 Radiation induced PCA components and PCA scores for the MCF7, PC3 and LNCaP cell lines 164
Figure 7.5 Averaged spectra from 20 unirradiated cells for the seven human tumour cell lines and the TRAMP-C2 mouse prostate tumour cell line 174
Figure 7.6 First three PCA components and PCA scores for the *in vitro* TRAMP-C2 irradiation experiment 176

Figure 7.7 Cell cycle PCA components and radiation induced PCA components for the TRAMP-C2 mouse prostate tumour cell line and the seven human tumour cell lines 178
Figure 7.8 Averaged spectra from 20 unirradiated and 20 irradiated cells from the TRAMP-C2 tumours grown and irradiated *in vivo*, with a pair of averaged Raman spectra from the *in vitro* TRAMP-C2 experiment 181
Figure 7.9 Cell cycle and radiation induced PCA components and PCA scores for the *in vivo* TRAMP-C2 irradiation experiment 183
Figure 7.10First three PCA components and PCA scores for the combined *in vitro* and *in vivo* TRAMP-C2 irradiation experiments data set 186

List of Abbreviations

3DCRT - 3D conformal radiation therapy
A - Adenine
A1450 - Area under the 1450 cm^{-1} peak
ATCC - American Type Culture Collection
BCCA-VIC - British Columbia Cancer Agency - Vancouver Island Centre
BL - Baseline
C - Cytosine
CARS - Coherent anti-Stokes Raman spectroscopy
CCD - Charge-coupled device
CR - Cosmic ray
DMEM - Dulbecco's Modified Eagle Medium
DNA - Deoxyribonucleic acid
DOF - Degrees of freedom
DRC - Deeley Research Centre
DSB - Double strand break
EBRT - External beam radiation therapy
FBS - Fetal bovine serum
FT - Fourier transform
FTIRM - Fourier transform infrared microscopy
G - Guanine
G0 - Gap 0 (phase)
G1 - Gap 1 (phase)
G2 - Gap 2 (phase)
HR - Homologous recombination
HWN - High-wavenumber
I - Interphase
IR - Infrared
IMRT - Intensity modulated radiation therapy
KHD - Kramer Heisenberg Dirac (expression)
LDA - Linear discriminant analysis
LP - Low-pass
LWN - Low-wavenumber

M - Mitosis
MLC - Multi-leaf collimator
MRI - Magnetic resonance imaging
MRS - Magnetic resonance spectroscopy
mt - mutant
NA - Numerical aperture
NHEJ - Non-homologous end-joining
PBS - Phosphate buffered saline
PCA - Principal component analysis
PET - Positron emission tomography
PI - Propidium iodide
RCIT - Radiation Cancer and Immunotherapy Theme
RM - Raman microscopy
RNA - Ribonucleic acid
ROI - Region of interest
RPMI - Roswell Park Memorial Institute (medium)
RS - Raman spectroscopy
S - DNA synthesis (phase)
SERS - Surface enhanced Raman spectroscopy
SF_2 - Surviving fraction after 2 Gy
SG - Savitsky-Golay
SNR - Signal-to-noise ratio
SRM - Signal removal method
SSB - Single strand break
T - Thymine
TA - Total area
TPLI - Three-point linearly interpolated
TPMEM - Two-point maximum entropy method
U - Uracil
UV - Ultraviolet
UVRRS - Ultraviolet resonance Raman spectroscopy
VMAT - Volumetric arc therapy
W - Window size
wt - wild-type

ACKNOWLEDGEMENTS

I would like to thank:

My family, for their unmeasurable support throughout my 5 years of graduate studies. A particular thank you to my wife, Jenn, for working so hard to provide me with the time and resources necessary to complete this project, and simultaneously being a fantastic partner and an amazing mother to our daughter. Special thanks also to my parents, whose time spent helping with childcare has made this thesis possible.

Dr. Andrew Jirasek, my supervisor, for mentoring, guidance, support, commitment, flexibility and understanding throughout this project.

My graduate committee, for support, encouragement, and valuable feedback and guidance throughout my studies.

National Science and Engineering Research Council, for providing me with four years of graduate scholarships.

I would also like to thank the staff of the Deeley Research Centre at the BC Cancer Agency's Vancouver Island Centre for providing the initial cell stocks, workspace, access to facilities and equipment, and technical assistance with cell culture and flow cytometry. Also, thank you to the physics staff at the BC Cancer Agency's Vancouver Island Centre for assistance with radiation delivery.

Chapter 1

Introduction

This thesis presents a body of work that exists at the intersection of radiation therapy physics, cellular radiobiology, and molecular spectroscopy. The results of this work pertain to the development and application of Raman spectroscopic techniques for analysis of radiation induced biochemical responses in human cancer cells. The motivations for this work are grounded in the aim to further understand the biochemical processes related to the interactions of ionizing radiation with human cells and tissues, and the hope to improve the medical practice of treating cancer with ionizing radiation. This first chapter provides an introduction to the fields of radiation therapy (section 1.1), cellular radiobiology (section 1.2) and biological Raman spectroscopy (section 1.3), in order to lay sufficient groundwork and justification for the remainder of the thesis work presented.

1.1 Radiation therapy

1.1.1 History

Medical practitioners have employed ionizing radiation for therapeutic and diagnostic purposes ever since Wilhelm Conrad Röntgen's discovery of the X-ray in 1895 [1, 2]. The first therapeutic use of radiation for treatment of superficial cancers occurred shortly after Pierre and Marie Curie's discovery and isolation of the radioactive elements polonium and radium in 1898 [3]. Decades later, the 1950s brought the development and implementation of cobalt treatment units and clinical linear accelerators [4]. This revolutionary advancement provided the means for clinical production of higher energy photons (greater than 1 MeV) allowing for increased penetration depth

in tissues for the treatment of deep-seated tumour sites. Further advancements in radiation delivery and treatment planning have led to modern treatment techniques such as 3D conformal radiation therapy (3DCRT) [5], intensity modulated radiation therapy (IMRT) [6], image-guided radiation therapy (IGRT) [7], tomotherapy [8] and volumetric arc therapy (VMAT) [9]. Brachytherapy, a term describing radiation therapy where a radiation source is placed inside or adjacent to the diseased site, has also underwent significant advancements due to developments in modern technology and treatment planning software, and is currently used in the treatment of many different types of cancers [10].

1.1.2 Modern clinical implementation

Radiation therapy is now the prescribed method of treatment for approximately one-third of all new cancer patients. It is estimated that over 50,000 Canadians will undergo radiation therapy for newly diagnosed cancer in 2011 [11]. Many patients respond successfully to radiation treatment, experiencing dramatic improvements in their quality of life and increasing their life expectancy. Some patients have much more limited levels of success. The outcome of a treatment may vary depending on the type and location of the tumour, the level of progression of the disease, and the individual response of a patient to the radiation treatment [12]. Almost all patients experience negative side-effects from radiation therapy, ranging from mild to debilitating, due to the unavoidable irradiation of healthy tissues [13]. As such, modern radiation therapy aims to deliver a lethal amount of radiation dose to the desired tumour volume, as prescribed by the radiation oncologist, while minimizing the irradiation of surrounding organs and tissues.

With modern medical imaging techniques (*e.g.*, ultrasound, computed tomography, magnetic resonance imaging, positron emission tomography) the position and shape of the tumour can be accurately determined and digitized, provided that the bulk of the tumour is easy to visualize. Areas of microscopic disease are often difficult to locate, and margins are added to the identified tumour bulk volume to hopefully encompass any microscopic extent [10]. Using modern radiation delivery techniques, such as 3DCRT, IMRT, VMAT or brachytherapy, treatments can be designed to conform to the 3D tumour volume, minimizing the dose to surrounding healthy tissue. Ideally, the radiation dose prescribed by the radiation oncologist for curative treatment of the tumour will be delivered to the desired location with minimal irradiation

of surrounding organs and tissues. The details of the various methods for the planning and verification of such clinical radiation treatments are described in detail in medical physics and radiotherapy texts [10, 14].

1.1.3 External beam radiation therapy

External beam radiation therapy (EBRT) is the delivery of high energy ionizing radiation to diseased tissues, where the radiation beam originates from outside of the patient and is targeted at the tumour. The most common method of delivering EBRT is with a linear accelerator, or "linac". The basic operation of a typical linac, and a brief description of the physical processes involved in depositing radiation dose, is presented below. A more thorough discussion of these concepts is available elsewhere in medical physics and radiotherapy texts [10, 14].

The clinical linear accelerator

A schematic of a typical modern linac design is shown in figure 1.1. An electron gun emits electrons into a waveguide, where the electrons are accelerated by ∼3 GHz microwaves to energies ranging from 4 to 24 MeV [10]. The electrons exit the waveguide and are bent by a 270° bending magnet to strike a metal target made of a high atomic number material, typically tungsten. The target stops the electrons and produces a high-energy photon beam via the bremsstrahlung interaction between the electrons and the atomic nuclei [14]. It should be noted that most clinical linacs can also deliver a therapeutic electron beam by replacing the target with an electron scattering foil [10], but this mode of treatment is less common than photon irradiation and is not further discussed here.

After exiting the target the lateral intensity distribution of the photon beam is forward-peaked, and is therefore passed through a flattening filter to produce a photon beam of uniform lateral intensity. The lateral geometry of the photon beam is initially determined by the primary collimators directly below the target (figure 1.1), and is further defined by two sets of moving paired collimator jaws (upper and lower) and an array of multi-leaf collimators (MLCs) positioned at the exit port of the linac treatment head. MLCs consist of ∼100 opposing pairs of individually controlled collimating filaments, typically 0.5 - 1.0 cm wide, allowing for variable and dynamic beam shaping. The inherent rate of the photon output of the linac, before the variable beam shaping by the jaws and the MLCs, is monitored by two ionization chambers

positioned below the flattening filter. The linac gantry can rotate 360° about the patient treatment couch, thus delivering radiation from any angle. The intersection of the gantry rotation axis and the beam axis is called the linac "isocentre", and is typically 100 cm from the source of the photon beam (*i.e.*, the bottom of the target). The rectangular field size of the photon beam at isocentre, as defined by the jaws, ranges from ∼1 cm x 1 cm up to ∼40 cm x 40 cm, with any further beam shaping provided by the MLCs.

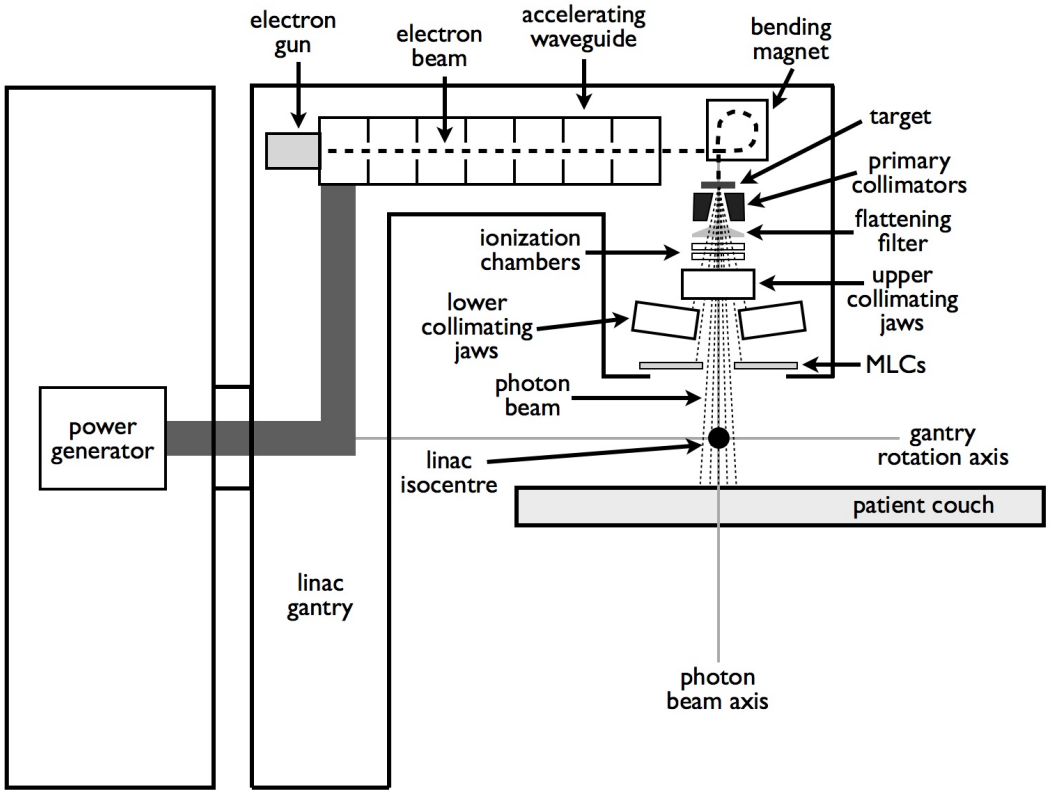


Figure 1.1: Schematic of a modern clinical linear accelerator for external beam radiation therapy.

Photon interactions in tissue

The photon beam exiting the linac treatment head has a broad energy spectrum, with energies ranging from ∼100 keV up to the maximum energy of the electrons striking the target (*i.e.*, 4 to 24 MeV). Photons at these energies predominantly interact with atoms in tissue through photoelectric, Compton, and pair and triplet production processes. The likelihood of each process occurring depends on the incident photon

energy and the effective atomic number of the material [14]. In terms of photon interactions at therapeutic energy levels, human soft tissue is approximately equivalent to water. The relative importance of each type of interaction, in terms of the relative number of interactions in water, is provided in table 1.1.

Table 1.1: Relative number of interactions (%) of photoelectric (τ), Compton (σ) and pair and triplet production (π) processes in water, for different incident photon energies [14].

Photon Energy (MeV)	τ	σ	π
0.010	95	5	0
0.026	50	50	0
0.050	11	89	0
0.100	1	99	0
0.150	0	100	0
1.00	0	100	0
2.00	0	99	1
4.00	0	94	6
10.00	0	77	23
24.00	0	50	50
50.00	0	29	71

Photoelectric interactions occur when an incident photon is absorbed by an atomic (inner shell) electron, causing the ejection of a photoelectron with energy equal to the difference between the incident photon energy and the binding energy of the electron. The electron vacancy in the inner shell causes an electron cascade as electrons from higher energy levels fill in the lower level vacancies, resulting in the emission of characteristic x-rays and Auger electrons. For incident photons with low energies ($\sim$10-20 keV), the photoelectrons are ejected at approximately right angles with respect to the incident photon direction, but the angular distribution of photoelectrons becomes more forward peaked (*i.e.*, $< 30^\circ$ from the direction of the incident photon) at energies of 1 MeV or higher [14]. In water, photoelectric processes only contribute significantly for photon energies up to $\sim$100 keV (table 1.1).

Compton scattering occurs when an incident photon scatters off a free (or valence shell) electron. The incident photon imparts kinetic energy to the scattered Compton electron equal to the difference between the initial photon energy and the scattered photon energy. The angular distribution of Compton electrons depends on the incident photon energy, and at $\sim$100 keV is fairly isotropic between 0° and 90° but becomes highly forward peaked (*i.e.*, $< 10^\circ$ from the direction of the incident photon) at energies of 10 MeV or higher [14]. In water, Compton scattering is the dominant

photon interaction for photon energies ranging from ∼50 keV up to ∼10 MeV (table 1.1).

Pair production occurs when a photon of energy greater than 1.022 MeV (twice the rest electron energy, 0.511 MeV) interacts with the atomic nucleus to form an electron-positron pair. The electron and positron each have kinetic energies equal to one-half the difference between the incident photon energy and 1.022 MeV (the energy required to create the electron-positron pair). The angular distribution of the ejected electrons and positrons becomes highly forward peaked as the incident photon energy increases [14]. The ejected positron travels through the material until its energy is reduced to approximately that of the surrounding electrons, upon which it annihilates with an electron to produce two photons emitted in opposite directions, each with energy 0.511 MeV. These annihilation photons then interact via photoelectric or Compton processes, as described above. Triplet production also occurs at high photon energies, but requires the initial photon energy to be greater than 2.044 MeV. Triplet production occurs when a photon interacts with a free electron to produce a positron-electron pair and an additional electron. Pair and triplet production processes contribute significantly to photon interactions in water at energies above ∼2 MeV (table 1.1), with triplet production occurring ∼10 times less often than pair production [14].

From a therapeutic perspective, the end result of the above photon interactions is the release of high-energy electrons (and positrons) with a forward peaked angular distribution, which interact with the atoms and molecules in the tissue. Electrons (and positrons) will interact with atoms either by (1) excitation, (2) ionization, (3) bremsstrahlung production, or (4) characteristic radiation production (via ejection of an inner shell electron) [14]. Processes (3) and (4) result in the creation of more photons (as will positron annihilation) that may undergo further interactions to release more electrons. Atomic excitations and ionizations, however, are processes where energy is absorbed locally by the tissue, possibly resulting in the breaking of molecular bonds, generation of reactive species, and biological damage. The amount of energy absorbed by a material, per unit mass, is defined as the radiation "dose" deposited in the material. Dose is typically expressed in units of Gray (Gy), where 1 Gy = 1 Joule / kg [14]. A further discussion of the effects of such energy deposition on biomolecules and cells is provided below in sections 1.2.2 and 1.2.3.

1.2 Cellular radiobiology

The field of radiobiology in medical physics attempts to understand the effects of ionizing radiation on cells and tissues in order to improve the effectiveness of radiation therapy treatments by incorporating knowledge of radiobiological processes, for tumours or healthy tissues, into the treatment planning process. Increased understanding of the fundamental interactions of radiation with biological tissues, and the radiation induced processes that determine cell death or survival post-irradiation, may potentially provide considerable benefits to both the individual patient and the field of radiation therapy. As discussed below, there are many outstanding problems and questions in the field of cellular radiobiology that currently limit the ability to further personalize and optimize modern radiation therapy treatments by incorporating knowledge of radiobiological effects. Before further discussion in this matter, the basics of the field of radiobiology, as applied to radiation therapy of human cells and tissues, is presented below.

1.2.1 The human cell: Biomolecules, cell cycle and genes

The cell is a collection of biological molecules organized into a single living unit. Human cells consist of membrane bound organelles suspended in cytoplasm, a saline solution that is 70-85% water, surrounded by a phospholipid bilayer membrane. Cell biomolecules are classified as either nucleic acids (DNA and RNA), proteins (*e.g.*, enzymes, structural or transport molecules), lipids (*e.g.*, fatty acids, phospholipids, cholesterol), or carbohydrates (*e.g.*, sugars, starch, glycogen). In terms of biomolecule composition, a cell is $\sim$79% protein, $\sim$5% nucleic acid, $\sim$11% lipid and $\sim$5% carbohydrate [15]. The results presented in this work are largely focused on changes in cellular concentrations of proteins, nucleic acids and lipids, with respect to cell cycle and genetic factors; as such, a brief discussion of each of these topics is provided.

Proteins are large molecules composed of a sequence of amino acids linked together by peptide bonds. There are 20 standard amino acids used for protein synthesis; the amino acid sequence is specified by the cell's genetic code and determines the type of protein synthesized. A single amino acid has an amine group, a carboxyl group, and a side-chain (**R**) specific to the amino acid (figure 1.2a). Amino acids join together via a dehydration reaction, forming an amide group (N-C=O) containing the peptide bond (C-N) and a water molecule (figure 1.2b). Long chains of amino acids arrange themselves into secondary protein structures, consisting of α-helices,

β-sheets and random coils, which further combine to form three dimensional tertiary and quaternary protein structures. The secondary structure of crambin, a small plant seed protein [16], is shown in figure 1.2c, with α-helices in red, β-sheets in blue and random coils in grey.

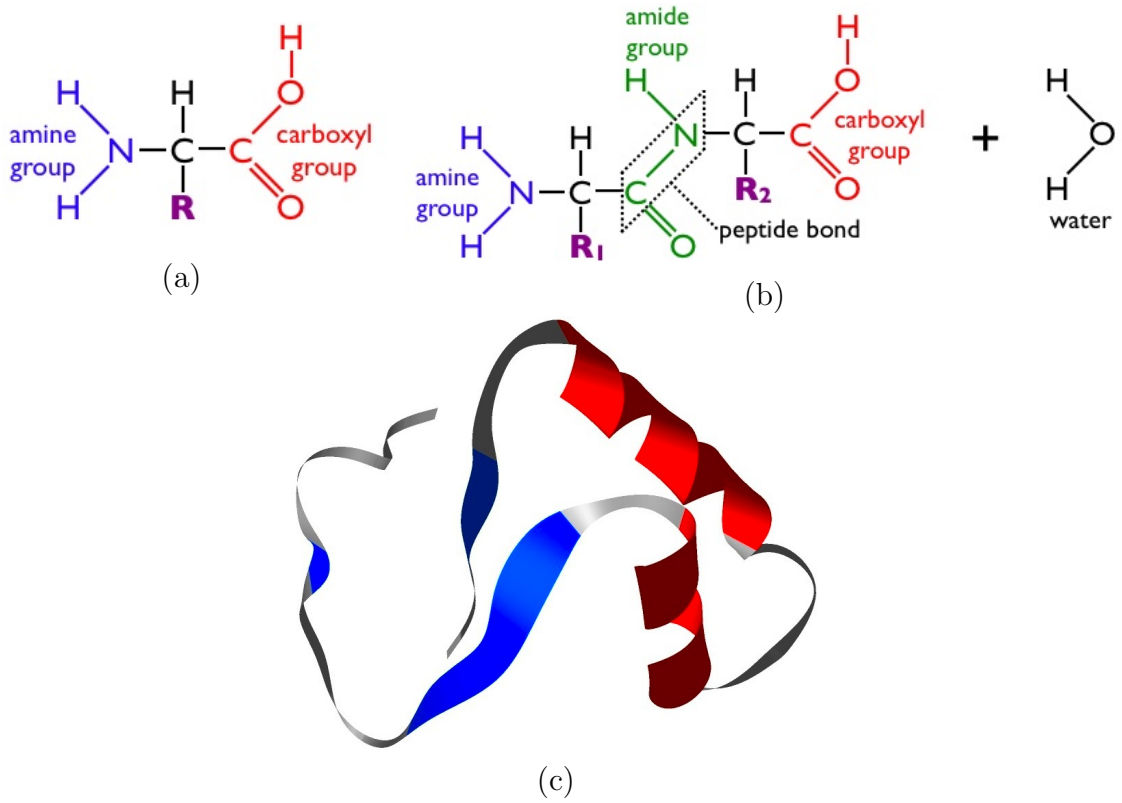


Figure 1.2: Schematic diagrams of (a) the general structure of an amino acid and (b) the joining of two amino acids in a dehydration reaction forming an amide group and a water molecule. (c) Secondary structure of the protein crambin [16], with α-helices in red, β-sheets in blue and random coils in grey (created using the "iMol Molecular Visualization Program", Piotr Rotkiewicz, 2007).

Nucleic acids in the cell consist of deoxyribonucleic acid (DNA) and ribonucleic acid (RNA). A single strand of a nucleic acid contains a chain of alternating sugar and phosphate groups which forms the backbone of the strand. Attached to each sugar group is a nitrogen-containing molecule called a base. The bases are classified into two groups based on their chemical structure: the pyrimidines thymine (T), cytosine (C) and uracil (U) have a single ring, and the purines adenine (A) and guanine (G) have a double ring. In double-stranded DNA, the bases on one strand form hydrogen bonds with bases on the other strand, and the two strands intertwine into a double-helix. Due to the chemical structure of the bases, T preferentially forms hydrogen bonds

with A, and C with G, making the two strands in the DNA helix complementary to each other (figure 1.3). RNA is structurally different from DNA: RNA molecules are primarily single-stranded and much shorter than DNA, the sugar group in the RNA backbone is ribose instead of deoxyribose, and the pyrimidine uracil replaces thymine as the complementary base to adenine (figure 1.3). The sequence of bases in DNA is the recipe for protein synthesis, the origin of genetic traits, and is responsible for programming the activities of the cell, whereas RNA is the corresponding machinery that facilitates protein synthesis and genetic expression through transcription and translation of DNA code.

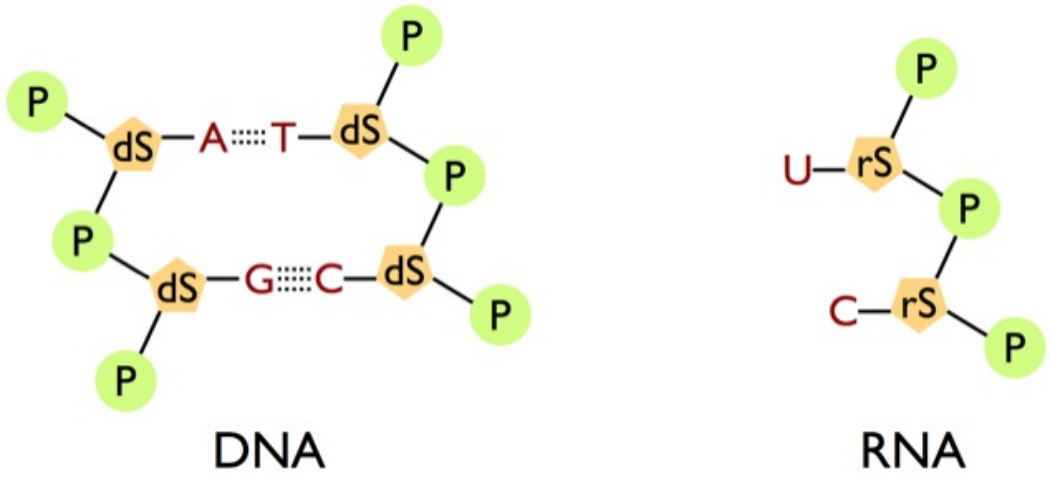


Figure 1.3: Block diagrams of a section of double-stranded DNA, showing the backbone of alternating deoxyribose sugar (dS) and phosphate (P) groups and the hydrogen bonding (dotted lines) of the bases adenine (A), thymine (T), guanine (G) and cytosine (C), and a section of single-stranded RNA, with ribose sugar (rS) and uracil (U) in place of thymine.

Lipids are a large and diverse class of biomolecules, whose biochemical roles in the cell include membrane structure, energy storage, cell signalling, biomolecule transport, and metabolism. Most lipids contain at least one fatty acid, which contains a long hydrocarbon chain typically 16-20 carbon units long. An example of a common cell membrane phospholipid is phosphatidylcholine (figure 1.4), which consists of a choline head group bonded with a phosphate group and a diglyceride (a glycerol molecule bonded to two fatty acids) with hydrocarbon chains 16 and 18 carbon units long.

The human mitotic cell cycle (figure 1.5) is divided into two main phases: interphase (I), a period of rest, growth and preparation for division, and mitosis (M), a

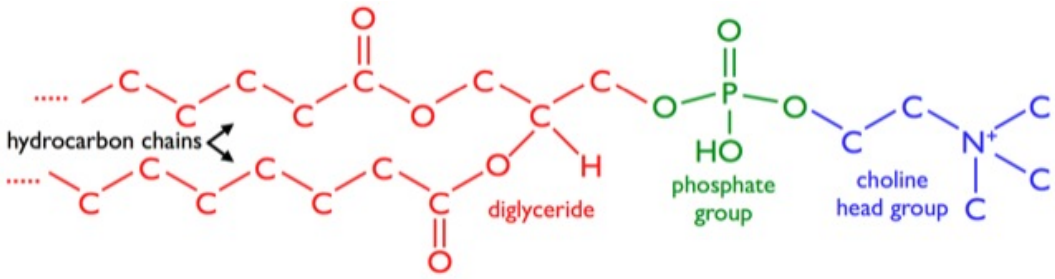


Figure 1.4: Schematic diagram of phosphatidylcholine, a cell membrane lipid consisting of a choline head group, a phosphate group, and a diglyceride. Most C-H bonds are omitted for clarity.

period of active cell division in which the cell is split into two daughter cells. The cell spends ∼50% (or more) of its cycle in interphase, which is further divided into three distinct phases: Gap 1 (G1) phase, DNA Synthesis (S) phase, and Gap 2 (G2) phase. The longest of these phases is G1 phase, in which the cell is growing in size and synthesizing proteins, lipids and RNA in preparation for DNA replication. Before onset of S phase, the DNA helices are packed with nuclear proteins into dense units called chromosomes. S phase begins with the onset of DNA replication and ends when the DNA has been fully copied. During DNA replication each chromosome is duplicated to form two sister chromatids. The sequence of bases is copied and shared between the two daughter cells during mitosis, thus providing each daughter cell with the necessary complement of DNA and passing genetic information to future generations. In G2 phase, the cell continues to grow and synthesize proteins, lipids and RNA in preparation for mitosis. It is also possible for a cell to suspend cell cycle progression and enter into Gap 0 (G0) phase, a state of cellular quiescence. G0 phase is a long-lived dormant phase, characterized by low RNA and protein levels [17, 18], which is usually entered into from early G1 phase.

Analysis of the cell cycle distribution of a population of cells is typically performed by measuring the relative DNA content of a representative sample of cells, using a technique called flow cytometry [19]. With this method, cells in suspension are treated with a fluorescent dye that binds to DNA. The cell suspension is then passed through a single-cell wide capillary and illuminated with an excitation laser. A photo-multiplier tube records the fluorescent intensity from each cell, which is proportional to the DNA content in each cell. In this fashion, G1, S and G2 phase cells can be distinguished based on their relative DNA content: G2 phase cells will have twice the amount of DNA as G1 cells, and S phase cells will have an intermediate amount of DNA relative

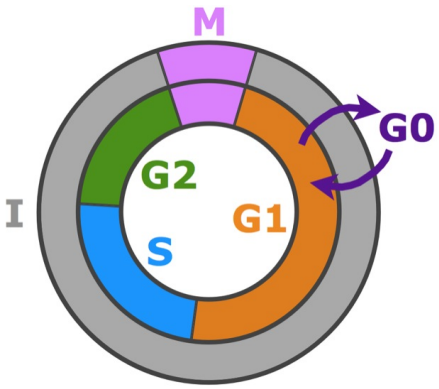


Figure 1.5: Phases of the mitotic cell cycle. The size of the region shown approximates the relative time for the corresponding phase to complete. I - Interphase, M - Mitosis, G1 - Gap 1, S - DNA Synthesis, G2 - Gap 2, G0 - Gap 0 (cellular quiescence).

to G1 and G2 phase cells. With modern flow cytometry technology, thousands of cells per second can be analyzed, enabling very fast and accurate measurements of the cell cycle distribution of a sample.

Distinct sequences of bases on chromosomes that code for specific proteins or nucleic acids are called genes. The order of the bases in a gene determines the order of the molecular subunits used to synthesize a protein or nucleic acid, thus determining the structure and function of the synthesis products to be used by the cell. The act of using the information coded in a gene to synthesize a functional product (*i.e.*, a protein or functional RNA molecule) is called gene expression, a vital act required for all normal cellular processes and for any active cellular responses to external stimuli. This role of DNA in the general health of a cell, and the ability of a cell to reproduce or adapt to its environment, has made DNA a primary biomolecule of interest in radiobiology.

1.2.2 Radiation effects on biomolecules

Photon and electron radiation can affect biomolecules in two ways. As discussed in section 1.1.3, photons of therapeutic energies interact predominantly through photoelectric, Compton and pair and triplet production processes, causing the release of high-energy (few keV to MeV) electrons. Direct interaction of these electrons with a biomolecule may cause atoms to be ionized or excited, causing damage to the biomolecule. This is called direct action of radiation [20]. More commonly, photons interact with components adjacent to a biomolecule (water in particular) and create

long-lived reactive species called free radicals. Free radicals can diffuse long distances (on the order of tens of nanometers [21]) compared to the width of a DNA strand ($\sim$2 nm), and react with and damage biomolecules due to the presence of unpaired electrons. This is called indirect action of radiation, and is the dominant source of damage from irradiation with photons and electrons (as opposed to irradiation with heavier particles such as protons, neutrons or α-particles) [20].

The creation of free radicals from water occurs when the absorption of radiation by water molecules results in a water ion pair (H_2O^+, H_2O^-) by the reactions

$$\begin{aligned} H_2O &\rightarrow H_2O^+ + e^- \\ H_2O + e^- &\rightarrow H_2O^- \end{aligned} \tag{1.1}$$

The water ions are unstable and rapidly dissociate in the presence of other water molecules by the reactions

$$\begin{aligned} H_2O^+ &\rightarrow H^+ + OH^\bullet \\ H_2O^- &\rightarrow OH^- + H^\bullet \end{aligned} \tag{1.2}$$

forming an ion pair (H^+, OH^-) and the free radicals $H^\bullet$ and $OH^\bullet$. The ion pair will likely recombine to a water molecule, but the free radicals have enough energy to diffuse distances great enough (*i.e.*, greater than the diameter of the DNA helix) to react with target biomolecules. It is estimated that about two-thirds of all indirect cell damage is caused by the highly reactive $OH^\bullet$ free radical [15, 20].

DNA is currently considered the primary cellular target of radiation, meaning that damage to DNA results in the highest probability of cell death [20]. Other biomolecules show radiation induced damage as well, such as chain breakage in carbohydrates, structural changes in proteins, and changes to the activity of enzymes [15]. The importance of these effects to the health of a cell is currently not well understood, and other biomolecules are not considered the primary target of radiation. As such, most studies in radiobiology have been focused on DNA damage and chromosomal aberrations resulting from radiation.

Radiation induced damage to a double-stranded DNA helix can manifest as a single-strand break (SSB) or a double-strand break (DSB), depending on the amount and spatial distribution of energy deposited by charged particles or free radicals at the site of interaction. SSBs are characterized by damage to a single base-backbone

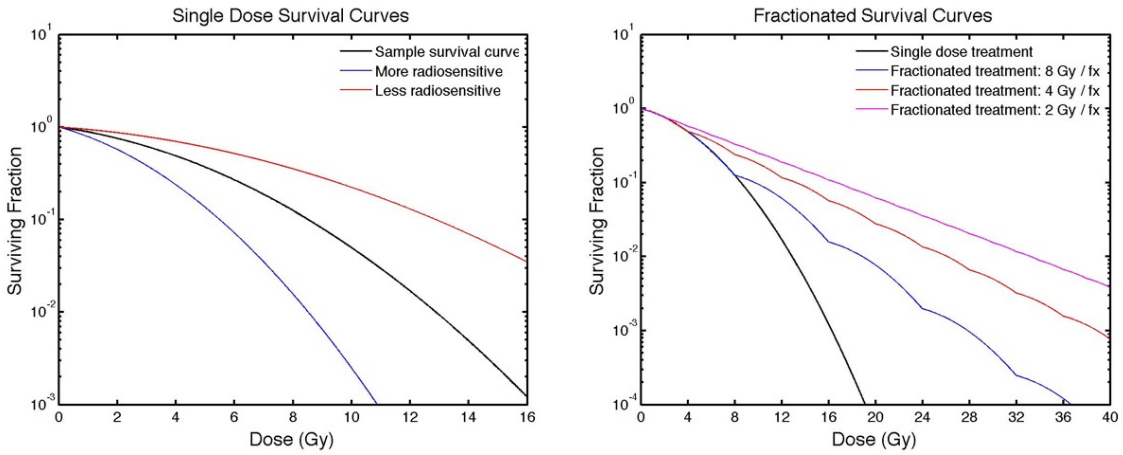
section, and are often repaired by the cell using the opposite strand as the complementary template. DSBs, however, result in a complete snapping of the DNA molecule. Repair of DNA DSBs is possible via one of two dedicated DSB repair mechanisms: non-homologous end-joining (NHEJ) or homologous recombination (HR) [20]. DNA DSBs can result in aberrations to an entire chromosome if not repaired, or if repaired incorrectly. There is a large body of work showing direct evidence of SSBs, DSBs, and both SSB and DSB repair in cells, using methods such as pulsed-field gel electrophoresis, single-cell electrophoresis, gas-chromatography mass-spectrometry, and fluorescence microscopy [22–26].

1.2.3 Radiation effects on cell survival

Chromosomal damage resulting from one or more DNA DSBs is likely the most lethal radiation effect in a cell, and is often directly visible with a high power microscope [20]. There has been significant work in this field, as the correct replication and splitting of the chromosomes during cell division may determine the ultimate health of the daughter cells. There are many different types of chromosomal aberrations, depending on the number of breaks in a chromosome or chromatid and depending on how the cell attempted to repair such breaks. Incorrect repairs can result in severely misshapen chromosomes, or chromosomes that appear normal but are coded incorrectly. Sufficient damage may result in the inability of the cell to divide properly, thus killing the cell in a process termed mitotic death. Alternatively, radiation damage may induce cell death independent of division via apoptosis, a controlled process of programmed cell death, or necrosis, an uncontrolled "violent" form of cell death. Damaged cells may also become senescent, a term describing cell inactivation by permanent cessation of cell cycle progression, although basic cellular processes (*e.g.*, metabolism, signalling) may continue for some time. Not all DNA or chromosome aberrations are lethal to the cell; some may instead cause a mutation (possibly carcinogenic) that does not affect the successful reproduction of the cell. Current radiobiological models employed to characterize the radiation response of cells are insensitive to the different types of cell death, mutations, or survival mechanisms, and focus purely on cell survival. These models are discussed below.

The effect of radiation on cells is commonly quantified experimentally with a cell survival curve, in which the surviving fraction of a population of irradiated cells is plotted against the delivered radiation dose, typically in units of Gray (Gy). The

surviving fraction is determined by irradiating a culture of cells to a known dose, sparsely seeding a known number of cells in a culture dish, incubating the cells to allow exponential growth of each surviving cell into a single colony (typically 1-3 weeks, depending on cell type), and then counting the number of colonies and dividing by the number of colonies formed by an unirradiated control culture [20, 27]. In conventional survival curves (figure 1.6a), the surviving fraction is plotted on a log scale, with dose on a linear scale. The curvature of the survival curve is determined by the radiosensitivity of the irradiated cells; a cell type that is more radiosensitive will require less dose to produce a given surviving fraction (figure 1.6a). For several reasons, such as accelerated growth rate and decreased repair rate, tumour cells are typically more radiosensitive than healthy cells of the same tissue type [15, 20]. This is one of the fundamental reasons why radiation therapy works for cancer treatment.



(a) (b)

Figure 1.6: (a) Single dose survival curves for mammalian cells with varying cell (or tissue) radiosensitivity and (b) fractionated dose survival curves (for constant radiosensitivity) with varying amounts of dose per fraction (fx). Reproduced from Quinn Matthews' M.Sc. thesis [28].

Experimental cell survival curves, plotted on a log-linear scale (figure 1.6a), are typically characterized by an initial linear portion at doses < 2 Gy (*i.e.*, the surviving fraction decreases exponentially with dose), followed by a downward curving portion at higher doses. At doses much higher than those delivered clinically for a single treatment fraction, experimental curves become linear again. The most widely accepted model for the shape of the survival curve is the linear-quadratic model, which assumes that there are two components contributing to cell death. The first component is linear with dose, corresponding to lethal DNA damage resulting from a single

electron, and the second component is quadratic with dose, corresponding to lethal DNA damage resulting from two independently created electrons. The form of the theoretical survival curve is therefore given by

$$S = e^{-\alpha D - \beta D^2} \tag{1.3}$$

where S is the surviving fraction, D is the radiation dose, and α and β are constants describing the relative contributions of the linear and quadratic components, respectively. For low doses (*i.e.*, $1 - 2.5$ Gy) that are typically prescribed for single daily treatment fractions, the model represents experimental data fairly accurately. However, at high doses, where experimental survival curves become linear, the curve described by the model bends continuously downwards as the quadratic term dominates, in disagreement with experiment.

In clinical radiotherapy practice, the delivery of a single high dose to a tumour is very rare. Instead, the total treatment is divided into fractions, in which multiple low dose treatments (typically $\sim$2 Gy each) are delivered, usually separated by $\sim$24 hours. For most forms of cancer treated with radiation there are significant advantages to fractionating a treatment. The delay between treatments allows healthy tissues to both repair any non-lethal damage and repopulate any lethally damaged regions with new healthy cells. This is doubly beneficial as the rate of repair is typically higher in normal tissues than in cancerous tissues. The delay also allows surviving tumour cells to progress to more radiosensitive phases of the cell cycle (primarily late G2 or M phase, in which rapidly dividing tumour cells inherently display a higher proportion than normal cells) prior to the next treatment. Furthermore, prolonging the treatment using multiple low doses helps to spare the patient from acute, early-onset side effects resulting from higher doses to healthy tissues. Therefore, through fractionation the total dose to the tumour can be significantly increased, thus improving the probability of eradicating the tumour. The effect of fractionated treatment on a cell survival curve is shown in figure 1.6b, indicating that as the dose per fraction is decreased, the fractionated curve increasingly approximates a continuation of the initial linear component of the curve. The linear-quadratic model described above is very useful for fractionated treatment planning, once the prescribed total dose has been determined [20].

1.2.4 Current problems and questions in radiobiology

In theory, the linear-quadratic model could be applied to experimental data to generate survival curves for various human tissues, tumours and organs, for the purposes of accurate treatment planning. However, most experimental radiobiology data is collected from established tumour cell lines cultured *in vitro* (*i.e.*, in a controlled environment in the lab), and it is often difficult to quantify a radiobiological effect for tissues *in vivo* (*i.e.*, in the original organism). Furthermore, it is difficult to obtain patient cell or tissue samples that will grow well enough in a lab to perform a radiation experiment and measure a survival curve [20]. There have been some successful *in situ* experiments, where irradiated cells are transplanted into mice and the tumour development is subsequently monitored, but there is some question on the validity of applying experimental observations from non-human organisms to human radiation therapy patients [29]. Due to these experimental difficulties, and a corresponding lack of understanding of cell death or survival probabilities with radiation for a given patient, current models are unable to predict the necessary doses required for complete control of a given tumour with acceptable levels of normal tissue damage. Prescribed doses are instead obtained by using population averages from past clinical treatments, determined from records of patient radiation response for healthy and cancerous tissues post-treatment, for a given dose and cancer type [12].

A major shortcoming of this method of dose prescription is the observed variability in the clinical response to radiation treatment between patients; both the level of normal tissue complications and the radiation response of the tumour are dependent on the individual sensitivity of the patient, or tumour, to radiation [30]. Some patients experience severe normal tissue complications that limit the amount of dose delivered, potentially compromising the effectiveness of the treatment [13]. Some tumours are more resistant to radiation compared to what is expected from the population average, which may lead to local recurrence of disease and consequently a very poor prognosis for further treatment [31]. As such, there is considerable interest in developing a predictive assay to determine individual radiosensitivity prior to, or during, treatment, such that the amount of dose delivered could be escalated for more resistant patients and reduced for more sensitive patients [30]. Experimental efforts to develop a predictive or monitoring assay for normal tissue complication probability have, so far, had unsatisfactory levels of success for clinical implementation [32–35]. Previous and current efforts to develop an assay to predict or monitor

tumour response are discussed below in section 1.2.5.

There are several other examples of poorly understood radiobiological phenomena that merit further investigation in the field of cellular radiobiology. As mentioned in section 1.2.3, the linear-quadratic model, used extensively for planning clinical radiation treatments, is relatively accurate at low doses used for fractionated treatments but significantly underestimates cell survival after delivery of single high doses [36]. Unknown radiobiological processes contributing to increased cell survival are likely the cause of this discrepancy. Furthermore, current models predict cell survival alone and are insensitive to radiation induced effects in cells that may alter the cell in some way, but may not be lethal, or even detrimental. An example of such an effect is the experimental observation of the cellular radioadaptive response [37], in which surviving cells exposed to low doses of radiation exhibit increased resistance to subsequent exposures at higher doses. This effect further indicates that radiation induced DNA damage alone does not determine the probability of cell death. Most cellular radiobiology experiments are only sensitive to DNA damage and/or repair, and do not monitor radiation induced effects on other types of biomolecules in a cell, namely RNA, proteins, lipids and carbohydrates [15].

The evidence of the radiobiological bystander effect is another important example of the need for further investigations across biomolecule types. First demonstrated in 1992 [38], it was seen that when a known fraction of cells in a culture is irradiated with low doses of α-particles, a higher fraction of cells show chromosomal damage than what was known to be irradiated. Many recent experiments have confirmed this effect. A common method is to use focused radiation microbeams, capable of irradiating only the nucleus of a single cell, to demonstrate that cells in the vicinity of the irradiated cell, but not exposed to the beam, display higher incidences of mortality, mutation and general damage [39]. Other experiments have harvested cells from an irradiated culture and transplanted them into a non-irradiated culture, subsequently observing the same damaging effects in the non-irradiated cells as in the irradiated cells, albeit to a lesser degree [40]. This is direct evidence of a radiation response that triggers some passage of information, or perhaps a toxic substance, from the damaged cells to the healthy cells. Such examples of radiobiological phenomena illustrate that there are indeed a number of unknown radiation induced biomolecule interactions that are vital to the radiation response of a cell culture, and that radiation effects determining cell death or survival are not restricted to DNA damage and repair.

1.2.5 Previous and current efforts for predicting or monitoring tumour response and intrinsic radiosensitivity

As mentioned above, there is considerable interest in developing an assay for predicting or monitoring tumour radiation response; however, there is currently no proven biochemical or imaging method for assessing tumour radiation response in a patient during the course of an extended treatment. An overview of the previous and current efforts in this field, and the important conclusions for future work, is provided below.

Studies using pretreatment tumour properties

Previous efforts to develop a predictive assay for tumour radiation response have targeted pretreatment tumour properties related to apoptosis [41], intrinsic tumour radiosensitivity [41–43], tumour oxygenation levels (hypoxia) [44–46] and tumour proliferation rate [47]. These previous efforts can be classified as either cell based assays, or functional assays.

Cell based assays, such as those investigating apoptosis or intrinsic tumour radiosensitivity, require the extraction of tumour cells for experimentation and/or analysis prior to radiation treatment. Apoptosis occurs as a cell death mechanism in unirradiated tumours and is often enhanced post-irradiation, raising the possibility that pre-treatment apoptosis levels may either predict tumour response to radiotherapy or be a surrogate marker for intrinsic radiosensitivity. As such, studies have been performed to measure the fraction of apoptotic cells in a tumour before treatment, and correlate values with treatment outcome. High levels of apoptosis were shown to be statistically predictive of decreased patient survival and increased local recurrence for cervical cancers, although apoptosis levels and intrinsic radiosensitivity were found to be uncorrelated [41]. However, subsequent studies investigating the predictive potential of pre-treatment apoptosis levels demonstrated variable levels of success [48]. Intrinsic tumour radiosensitivity can be measured directly by a clonogenic survival assay [27], provided that the extracted tumour cells will grow and survive long enough to perform a radiation experiment and measure a surviving fraction. Radiosensitivity is often quantified simply by the surviving fraction of a culture of tumour cells after 2 Gy of radiation (SF_2). Several studies have established that SF_2 is statistically predictive of local tumour control in radiotherapy patients, independent of other factors [41–43]. Unfortunately, not all patient tumours can be analyzed successfully with this technique (*e.g.*, 71% and 63% success rates reported for studies on cervical [42] and

head and neck [43] tumours, respectively) and the assay times are too long (∼4 weeks) for such assays to guide clinical radiation therapy treatments. However, the result that SF_2 is statistically predictive of local tumour control confirms that understanding the fundamental cellular mechanisms affecting intrinsic radiosensitivity is a vital component for improving the effectiveness of future radiation therapy treatments.

Functional assays directly measure some property of the tumour, such as oxygenation status or proliferation rate. Low oxygen concentration (hypoxia) in a tumour is known to cause increased tumour cell survival. This phenomenon is generally believed to be the result of a decrease in the "oxygen fixation" of radiation induced damage. Biomolecular damage caused by the $OH^{\bullet}$ radical can often be restored to its original undamaged form by reaction with a H_3O^+ ion, whereas the reaction of a damaged molecule with oxygen will chemically "fix" the damage, thus necessitating biological repair [31]. As such, many studies have investigated tumour hypoxia for the prediction of treatment outcome, using either direct measurements of tumour oxygenation levels using polarographic electrodes [44], or indirect measurements of hypoxia by the detection of certain proteins expressed by hypoxic cells [45]. However, direct hypoxia measurements are highly invasive and difficult to accurately reproduce, and both direct and indirect methods have shown variable levels of success for prognostic significance [46]. Functional assays measuring parameters related to tumour proliferation rate have also been investigated for predicting radiation response. Such assays are performed by injecting patients with a drug known to preferentially accumulate in rapidly dividing cells, and then later extracting tumour biopsies for analysis of drug uptake. However, such methods are inherently invasive due to the toxicity of the injected drug, and have shown only weak correlations between increased proliferation and poorer treatment outcome [47].

Although many cell based or functional assays have demonstrated predictive potential, most have had unsatisfactory levels of success when correlated with treatment outcome, or have not been reproduced by other studies using different techniques or studying different tumour sites, or have posed significant technical difficulties preventing clinical implementation [48].

Gene-based techniques for predicting intrinsic radiosensitivity

Due to mounting evidence [48], it is now widely accepted that genetic differences between tumours from different patients play a significant role in the observed differences

in intrinsic tumour radiosensitivity. As such, the relationship between the genetic status of tumours and intrinsic radiosensitivity has been studied extensively. However, the dependence of the functional state of individual genes (*e.g.*, p53, a gene known to regulate cell cycle progression and the initiation of apoptosis post-irradiation [49]) in determining radiosensitivity or radioresistance appears to depend on the tumour cell lines examined [49–53]. For example, one study, using a variety of radioresistant and radiosensitive cell lines derived from cancers of the brain, bladder and ovary, reported that functional p53 is a requirement for increased sensitivity to radiation [49]. However, a later study reported increased radiation survival in prostate cancer cells with functional p53 as compared to cells with non-functional mutated p53 [51]. These conflicting results suggest that many genes (or other factors) are likely involved in determining radiation response.

Recently developed methods have applied the expression profiles of multiple genes to predict the tumour radiosensitivity of a patient by comparisons with expression profiles and experimental radiation survival data from established tumour cell lines [54, 55]. These methods have been shown to be statistically predictive of tumour response in esophageal and rectal cancers, and of local tumour control in head and neck cancers [56]. Such techniques likely have the most potential for clinical implementation, as pre-treatment genetic profiling of a patient's tumour cells is relatively simple with modern techniques. However, both pre-clinical [54, 55] and clinical [56] studies using such techniques have reported many false positives and negatives, which may degrade confidence in such methods for guiding personalized treatments. Such methods may also be inherently limited by the use of laboratory data from a limited panel of established tumour cell lines upon which the models are constructed, possibly limiting the application for clinical cases across a variety of tumour types. Furthermore, there may be other sources of experimental bias in these studies; for example, none of the genes identified as being important for predicting intrinsic radiosensitivity are genes known to be expressed under hypoxic conditions (and therefore likely important in the tumour radiation response of a patient), as all radiation experiments used to generate the predictive model were performed under fully oxygenated conditions.

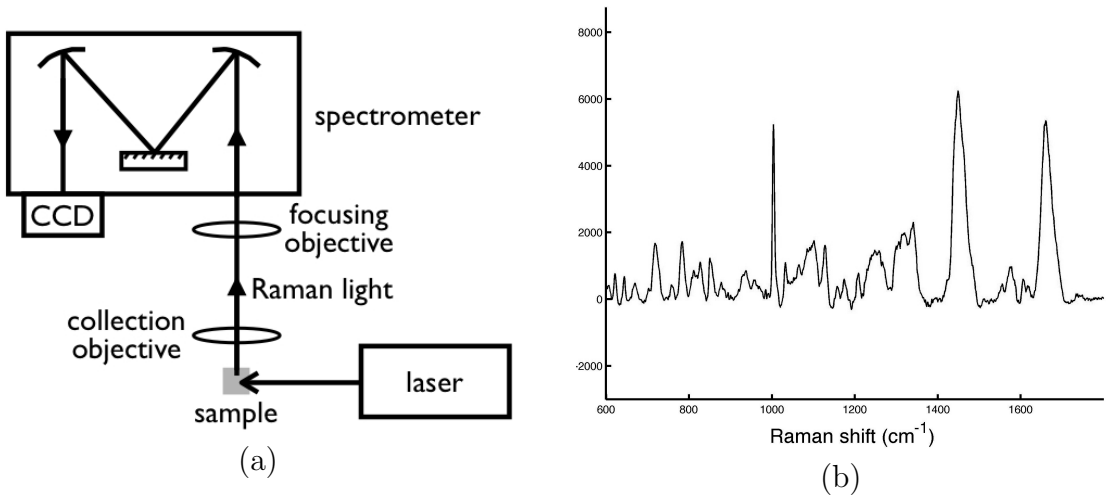
Need for new methodologies?

In light of these previous and ongoing research efforts, it is likely that future advances in the field of experimental radiobiology as applied to personalized radiation therapy

may benefit from the use of new biochemical analysis methods with the ability to analyze biochemical radiation response *in vitro* or *in vivo* across a wide variety of biomolecules. One such technique is Raman spectroscopy (RS). An brief introduction to RS, and a description of its previous applications in the biomedical sciences, is provided below.

1.3 Raman spectroscopy in cell and tissue analysis

Raman spectroscopy (RS) is a vibrational spectroscopy technique in which an optical wavelength laser is focused onto a sample, inducing transitions between molecular vibrational levels and creating inelastically scattered photons with frequencies and intensities characteristic of the molecules in the sample. A schematic of a general Raman system is shown in figure 1.7a. The scattered Raman photons are passed through a spectrometer and collected on a charge-coupled device (CCD) for spectroscopic analysis. The resulting Raman spectrum (*e.g*, figure 1.7b) provides a detailed description of the molecular composition within the sampling volume, as each peak can be assigned to the vibration of either a specific type of molecule or a molecular group from a specific class of molecules.



(a) (b)

Figure 1.7: (a) General schematic of the light path for a typical Raman spectroscopy system. (b) Sample Raman spectrum of a single cell.

A brief overview of the advantages of RS for biomedical applications is provided below (section 1.3.1). This is followed by a review of the previous RS studies of cells and tissues (section 1.3.2) and radiobiological processes (section 1.3.3), which are

directly relevant to the scope of this thesis (section 1.4). An overview of RS theory and implementation, and further examples of RS applications, is provided in chapter 2.

1.3.1 Advantages of Raman spectroscopy for biomedical applications

RS is attractive for biomedical applications for a number of reasons. Within a specific range of laser wavelength and power, RS is noninvasive and nondestructive, allowing the analysis of live cells or tissues without perturbation of the sample [57–59]. The use of high-power focusing optics can provide spatial resolutions as low as 1 μm, well below the typical size of a human cell (10-50 μm diameter). This sampling flexibility allows RS to be applied to a variety of biological samples ranging in scale from sub-cellular organelles to multi-cellular bulk tissues or histopathological tissue sections. For cellular RS analysis, it has been demonstrated that RS can provide biochemical information from live cells at comparable levels of accuracy and sensitivity as established techniques such as magnetic resonance spectroscopy (MRS) and flow cytometry [60]. Furthermore, RS has an inherent advantage over other biochemical analysis tools by enabling the simultaneous detection of a variety of molecular structures across proteins (*e.g.*, amino acids, conformational structures), nucleic acids (DNA and RNA) and lipids (*e.g.*, cholesterol, choline, CH_2 groups), in a single acquisition. This utility allows complex biochemical changes in cells to be analyzed simultaneously across different classes of biomolecules rather than analyzing a single type of molecule (*e.g.*, DNA), or class of molecules (*e.g.*, metabolites), an inherent restriction of many established techniques. In addition, RS is very sensitive to changes in molecular structure, especially the secondary structure of proteins [61]. RS also has the potential to be applied *in vivo* with the use of fiber-optic probe technology [62].

1.3.2 Raman spectroscopy of cells and tissues

The first successful application of RS for single-cell analysis was reported in 1990 [63]. One of the first biomedical applications of modern RS techniques was to investigate the spatial distribution of the molecular composition of bacteria and other medically relevant microorganisms [64–67]. In recent years, there has been considerable inter-

est in applying RS to live cells and tissues for cancer detection and diagnosis. Due to the complexity of a single biological Raman spectrum, RS is commonly used in conjunction with multivariate statistical methods such as principal component analysis (PCA), linear discriminant analysis (LDA) or cluster analysis. RS with the use of such multivariate methods has been successfully applied to discriminate between healthy and cancerous skin [68–70], bladder [71] and gastric [72, 73] tissues, and to aid in the histopathological analysis of prostate cancer [74]. Applying RS and multivariate methods to the analysis of single cells has shown the ability to distinguish between healthy and tumourogenic rat fibroblast cells [75], human bone cells [76] and human epithelial cells from a variety of organs [77]. Similar techniques have been employed to discriminate between different types of human tumour cell lines from a mixed sample set [78–80].

The single cell studies mentioned above have focused on the differences in Raman spectra between different cell types (*i.e.*, healthy vs. tumour, prostate vs. bladder). In addition, several studies have used RS for analyzing biochemical differences arising within a population of a single type of cell, which is also a research goal of this work. One pair of studies detected spectral changes from cell death via apoptosis [81] and spectral differences between live and dead cells [59], and another pair of studies detected spectral changes due to cell death via necrosis [82] and spectral differences between exponentially growing (proliferative) and plateau phase (non-proliferative) cells [83]. One recent study applied PCA and LDA to Raman spectra of single live cells to discriminate between cells synchronized in the G0/G1, S and G2/M phases of the cell cycle, demonstrating RS discrimination of G0/G1 cells from S and G2/M cells [84].

1.3.3 Raman spectroscopy in radiobiology

Conventional RS techniques, without single-cell spatial resolution, have already had some success in the field of radiobiology. Both Fourier-transform Raman spectroscopy and ultraviolet resonance Raman spectroscopy (UVRRS) have been used to investigate radiation induced spectral changes in isolated aqueous DNA [85, 86]. However, these studies report poor RS sensitivity to radiation damage, as achieving a measurable change in the spectra of isolated DNA using this modality requires very high doses (hundreds to thousands of Gray), which are well above the levels used in clinical radiation therapy and well above the levels known to cause over 99.9% mortality in

a typical *in vitro* cell culture (figure 1.6a). Other RS studies have detected radiation induced changes in sodium hyaluronate, an important molecular component of the extra-cellular matrix in tissues [87], and radiation induced changes in protein and lipid membranes [61, 88–91]. RS has also been used to investigate the effect of proton irradiation on both healthy and cancerous human skin samples, and was successful in demonstrating increased sensitivity to protein damage in cancerous tissues [92]. RS was also successful in measuring changes in protein levels in mice tissues after irradiation of the brain [93]. A recent clinical RS study successfully discriminated between responding and non-responding cervical cancers post-irradiation [94], yet made no conclusions regarding the differences in the biochemical composition between tissues or the molecular basis for RS discrimination. However, this study does validate the potential for using RS in clinical implementation.

These previous studies suggest that RS may be useful for investigating the radiation response of human cells and tissues. However, RS has not yet been applied to study the effects of ionizing radiation on single living cells, which are the primary experimental source from which radiobiological models are inferred [20]. Radiobiological RS experiments on single cells may shed new light on some of the problems and questions discussed in section 1.2.4. Additionally, RS investigations of human tumour cells may break new ground in the attempt to understand differences in tumour radiosensitivity, as discussed in section 1.2.5. Some attractive future clinical applications of RS techniques are the development of a predictive assay for the radiation response of healthy or tumour tissues, or the development of a method to assess radiation response during treatment.

1.4 Thesis scope

There are three primary goals of this work: (1) to develop a robust methodology for RS acquisition and data processing for highly sensitive and specific spectral analysis of biochemical changes in single human cells, (2) to demonstrate the utility of such RS techniques for analysis of radiation induced biochemical changes in a human tumour cell line, and (3) to apply RS to investigate the biochemical radiation response of a panel of human tumour cell lines with varying intrinsic radiosensitivity. The achievement of these three goals is presented in this work in four parts, each of which are introduced below.

The first part of this work (chapter 4) validates the development and implementa-

tion of a novel combination of RS acquisition and data processing techniques, coupled with principal component analysis (PCA), for biochemical analysis of single human cells. A representative set of raw data is used to demonstrate the effect of varying certain processing parameters on the final outcomes and interpretations drawn from the data. This analysis demonstrates the effectiveness and robustness of the developed techniques, and serves to identify the potential pitfalls that may arise during implementation of the required data processing steps.

The second part of this work (chapter 5) applies RS to investigate and characterize the inherent sources of spectral variability in a human prostate tumour cell line (DU145). The results of RS studies on both asynchronous and synchronous cell cycle distributions are presented, alongside cell cycle data from flow cytometry, to identify and characterize spectral variability arising from differences in cell cycle and *in vitro* culture conditions. These results provide the basis for interpretation of the subsequent studies presented in this work.

The third part of this work (chapter 6) demonstrates the RS observation of radiation induced spectral changes in a human tumour cell line (DU145). PCA is shown to effectively separate spectral variations arising from cell cycle differences (characterized in chapter 5) from those arising from a biochemical radiation response.

The fourth part of this work (chapter 7) extends the same radiation RS methods used for the DU145 cell line (chapter 6) to multiple human tumour cell lines derived from prostate (DU145, PC3, LNCaP and PacMet), breast (MCF7 and MDA-MB-231) and lung (H460). The cell lines differ by tissue of origin, genetic status and radiosensitivity. This study presents the observation of unique RS radiation response signatures that correlate with known radiosensitivity parameters (SF_2) and segregate according to p53 gene status, corroborated by cell cycle analysis post-irradiation. Again, PCA separates radiation induced spectral changes from other sources of spectral variability, such as cell cycle. Candidate radiation response mechanisms affecting radiation survival are discussed with respect to the biochemical changes observed with RS. Preliminary results obtained from a mouse prostate tumour cell line (TRAMP-C2), irradiated both *in vitro* and *in vivo* (as solid tumours in mice), are also presented and discussed.

Before the results of this work are presented, an overview is provided of RS theory and modern application (chapter 2). This is followed by a brief description of the materials and methods used throughout this work (chapter 3). The thesis concludes (chapter 8) with a summary of the key results of the work and a discussion of some

future efforts planned in this field. Such future efforts include the RS analysis of a larger panel of tumour cell lines, complementary biochemical experiments to elucidate the biological nature of the RS observations, and solid tumour irradiation experiments using human tumour cell lines grown in mice.

Chapter 2

Raman Spectroscopy

This chapter provides an overview of Raman spectroscopy (RS). A brief history of the development of RS is presented (section 2.1), followed by a theoretical description of Raman scattering and molecular vibrations (section 2.2). Next, an overview of the instrumentation for the application of RS and Raman microscopy (RM) is provided (section 2.3). Some important considerations that arise when performing RS on biological samples such as cells and tissues are then discussed (section 2.4). The chapter concludes with a discussion of the advantages and disadvantages of RS (section 2.5) and a description of several of the more prevalent modern applications of RS (section 2.6).

2.1 History

The "Raman effect" is the name given to the inelastic scattering of light of optical wavelengths. The initial theory was published in 1923 by A. Smekal [95]. The Raman effect was first observed experimentally in 1928 by C. V. Raman and K. S. Krishnan [96]. Raman and Krishnan used a telescope and a series of optical filters to isolate blue light from the sun, focus it onto a transparent liquid sample of carbon tetrachloride, and observe scattered green light [97]. Shortly after this experiment a mercury lamp was substituted as the incident light source and the apparatus was used to measure the scattered Raman spectrum of benzene [98]. Benzene was the first molecule to be analyzed in detail using Raman spectroscopy.

Following Raman's discovery, efforts were made to improve the excitation source. Various types of lamps were implemented, using elements such as helium, bismuth,

lead, zinc and mercury; mercury lamps proved the most effective due to their high intensity [99–101]. The development and improvement of mercury lamps for Raman excitation continued until the introduction of lasers in 1962 [102]. Lasers proved to be excellent Raman excitation sources and are still used in modern Raman applications due to their high power, ease of collimation and focusing, excellent monochromaticity, and availability in a range of wavelengths from the ultraviolet (UV) to the infrared (IR).

A similar technological evolution occurred in the improvement of detection methods for the Raman scattered light. The first measurements of Raman scattering were performed using photographic plates, but the need for long integration times and photographic plate development made this a highly inefficient method. The first photoelectric Raman systems were developed in the 1940's and 1950's using photomultiplier tubes for light collection, providing a significant increase in detection sensitivity [103, 104]. However, obtaining a multi-wavelength Raman spectrum with a single detector is highly inefficient, requiring the user to scan the desired frequency range and collect signal at each point. Multi-channel detection using photo-diode arrays was introduced in the early 1980's [105, 106], and was soon followed by detection using charge-coupled device (CCD) arrays [107, 108]. CCDs are still the detectors of choice in modern systems due to their low readout noise, high efficiency, and excellent sensitivity over a wide range of wavelengths.

The advancement of RS into its modern form has also been aided by technological improvements in diffraction gratings, monochromator and spectrometer design, optical filters and many other components. There are many types of RS techniques currently being utilized in both industry and research, such as Fourier transform (FT) RS, surface enhanced RS (SERS), fiber-optic RS, UV resonance RS (UVRRS) and coherent anti-Stokes RS (CARS). Each Raman variant has its own advantages and disadvantages, depending on the sample being analyzed, but all are currently used with great success in a variety of scientific disciplines and industrial applications.

2.2 Theory of Raman scattering

2.2.1 Origin of the Raman effect

If a molecule is illuminated with optical photons of frequency ν_0 and energy $h\nu_0$ (h is Planck's constant), most of the the light is elastically scattered with no frequency

shift (Rayleigh scattering). However, approximately one in every 10^6 photons [109] induces a transition between vibrational energy levels of the molecule. This results in an inelastically scattered photon with frequency $\nu' = \nu_0 \pm \nu_m$ (Raman scattering), where ν_m is the positive difference in vibrational frequency between the initial and final vibrational states of the molecule. The $\nu_0 - \nu_m$ frequency shift is known as "Stokes" scattering and occurs when the scattering photon loses energy $h\nu_m$ as the molecule is excited to a higher energy vibrational state. The $\nu_0 + \nu_m$ shift is known as "anti-Stokes" scattering and occurs when the scattering photon gains energy $h\nu_m$ as the molecule de-excites to a lower energy vibrational state.

In most Raman spectroscopy applications, the energy of the incident photons is chosen to be less than the energy of the first electronic excited state of the molecule or molecules of interest. As discussed further below, a molecule at room temperature most likely exists in its ground vibrational state. When illuminated, the molecule is excited from its ground state to a virtual excited state, a distorted electron distribution that is highly unstable, and decays almost immediately to either the original state, in the case of Rayleigh scattering, or an excited vibrational state, in the case of Stokes Raman scattering (figure 2.1). Anti-Stokes Raman scattering requires the molecule to initially exist in an excited vibrational state, and arises from the decay from a virtual excited state to the ground vibrational state (figure 2.1).

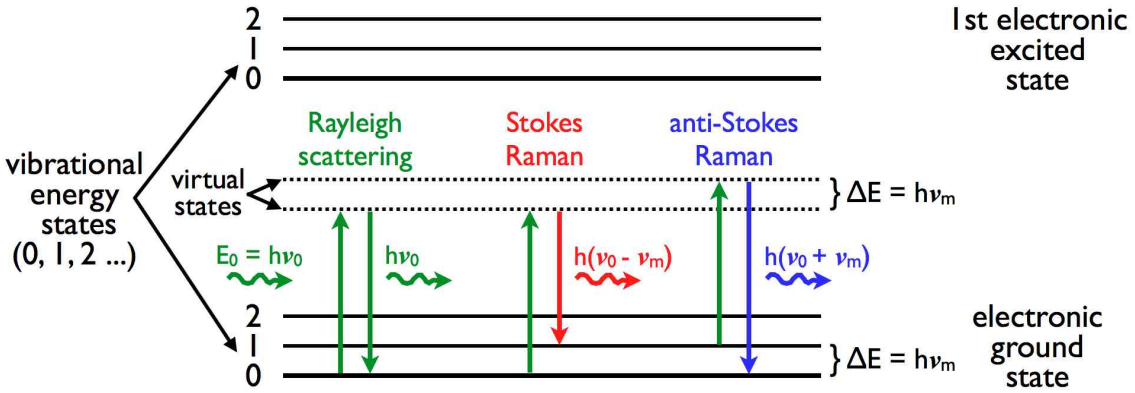


Figure 2.1: Energy level diagram for a molecule irradiated with optical photons of frequency ν_0 subject to either elastic Rayleigh scattering, or inelastic Stokes or anti-Stokes Raman scattering.

The origin of Raman scattering from a molecule can be derived using a classical model in which the illuminating laser source is expressed as an electric field vector $\mathbf{E}$

with frequency ν_0, oscillating with time as

$$\mathbf{E} = \mathbf{E_0} \cos 2\pi\nu_0 t \tag{2.1}$$

where $\mathbf{E}_0 = (E_{0x}, E_{0y}, E_{0z})$ is the electric field amplitude.

If a molecule is placed in the electric field of equation 2.1, the field induces a charge separation within the molecule that creates an electric dipole moment $\mathbf{P}$, given by

$$\mathbf{P} = \alpha\mathbf{E} \tag{2.2}$$

where the proportionality constant α is called the *polarizability* of the molecule. Expressed in matrix form with cartesian coordinates, α becomes a 3×3 polarizability tensor and equation 2.2 takes the form

$$\begin{bmatrix} P_x \\ P_y \\ P_z \end{bmatrix} = \begin{bmatrix} \alpha_{xx} & \alpha_{xy} & \alpha_{xz} \\ \alpha_{yx} & \alpha_{yy} & \alpha_{yz} \\ \alpha_{zx} & \alpha_{zy} & \alpha_{zz} \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix} \tag{2.3}$$

The polarizability tensor is symmetric for normal Raman scattering, and the molecular vibration will yield Raman scattered photons if any one of the polarizability tensor's components is changed during the vibration [110], as shown below.

If the molecular charge separation induced by the electric field causes vibrations of the atomic nuclei with frequency ν_m, then the displacement from the equilibrium position of the ith nucleus, q_i, is written as

$$q_i = q_{i0} \cos 2\pi\nu_m t \tag{2.4}$$

where q_{i0} is the vibrational amplitude of the ith nucleus. For small amplitude vibrations the polarizability α is linear with q_i and can be expanded about the nuclear equilibrium position as

$$\alpha = \alpha_0 + \left(\frac{\delta\alpha}{\delta q_i}\right)_0 q_i + \ldots \tag{2.5}$$

where α_0 is the equilibrium polarizability, and $(\delta\alpha/\delta q_i)_0$ is the rate of change of α with respect to q_i, evaluated at the equilibrium position.

If equations 2.1 and 2.5 are substituted into equation 2.2, the induced dipole

moment becomes

$$\begin{aligned} \mathbf{P} &= \alpha \mathbf{E_0} \cos 2\pi\nu_0 t \\ &= \alpha_0 \mathbf{E_0} \cos 2\pi\nu_0 t + \left(\frac{\delta\alpha}{\delta q_i}\right)_0 q_i \mathbf{E_0} \cos 2\pi\nu_0 t \end{aligned} \tag{2.6}$$

and using equation 2.4 for q_i, equation 2.6 becomes

$$\mathbf{P} = \alpha_0 \mathbf{E_0} \cos 2\pi\nu_0 t + \left(\frac{\delta\alpha}{\delta q_i}\right)_0 q_{i0} \mathbf{E_0} \cos(2\pi\nu_0 t) \cos(2\pi\nu_m t) \tag{2.7}$$

Finally, making use of the trigonometric identity $2\cos x \cos y = \cos(x-y)+\cos(x+y)$, equation 2.7 becomes

$$\mathbf{P} = \alpha_0 \mathbf{E_0} \cos 2\pi\nu_0 t + \frac{1}{2}\left(\frac{\delta\alpha}{\delta q_i}\right)_0 q_{i0} \mathbf{E_0} [\cos\{2\pi(\nu_0 - \nu_m)t\} + \cos\{2\pi(\nu_0 + \nu_m)t\}] \tag{2.8}$$

Classical electrodynamics predicts that the first term in equation 2.8 describes an oscillating dipole that radiates light of frequency ν_0, corresponding to Rayleigh scattering. Likewise, the second term describes the radiation of frequencies $\nu_0 - \nu_m$, corresponding to Stokes Raman scattering, and $\nu_0 + \nu_m$, corresponding to anti-Stokes Raman scattering. However, if $(\delta\alpha/\delta q_i)_0$ is zero (*i.e.*, the polarizability tensor (equation 2.3) is unchanged with displacement of the nuclei) the molecular vibration is not Raman-active and does not produce any Raman scattering. A discussion of Raman-active and Raman-inactive vibrations is provided in section 2.2.3.

A misleading aspect of equation 2.8 is that it predicts equal contributions from Stokes ($\nu_0 - \nu_m$) and anti-Stokes ($\nu_0 + \nu_m$) scattering. Experimental observations show that Stokes scattering is the dominant process. This is explained by the relative population of the vibrational states depicted in figure 2.1. The population ratio of an excited vibrational state (n) to a lower vibrational state (m) is determined by the Boltzmann distribution

$$\frac{P_n}{P_m} = \frac{g_n}{g_m} \cdot e^{-(E_n - E_m)/kT} \tag{2.9}$$

where $E_n - E_m$ is the energy difference between the two states, k is Boltzmann's constant, T is the absolute temperature, and g is the degeneracy of a vibrational level. Therefore, as the temperature increases or the energy gap between levels m and n decreases, the fraction of molecules in the higher vibrational energy state increases, leading to a larger anti-Stokes effect. However, for most biological molecules and

bond types ($E_1 - E_0 = 400 - 4000$ cm^{-1}) at room temperature ($kT \approx 200$ cm^{-1}) the population primarily exists in the ground vibrational state (P_n/P_m ratio between $\sim 1\times10^{-1}$ and $\sim 2\times10^{-9}$), leading to a much stronger Stokes contribution [110]. Most Raman applications, with the exception of coherent anti-Stokes Raman spectroscopy (CARS), exclusively measure Stokes scattering; likewise in this work the anti-Stokes contributions are not considered.

2.2.2 Intensity of Raman scattering

An expression for the intensity of Raman scattering requires a quantum mechanical approach, which accounts for the discretization of vibrational states and describes the relationship between molecular properties and Raman scattering [109]. The full details are found elsewhere [109–111], and only a brief discussion is presented here.

The intensity of Stokes scattered radiation for a transition from a vibrational state m to a vibrational state n of higher energy can be expressed as

$$I_{mn} = C \cdot I_0 \cdot (\nu_0 - \nu_{mn})^4 \cdot \sum_{\rho\sigma} |(\alpha_{\rho\sigma})_{mn}|^2 \tag{2.10}$$

where C is a constant, I_0 is the incident laser intensity, ν_{mn} is the frequency difference between vibrational states m and n, and $(\alpha_{\rho\sigma})_{mn}$ is the change in the polarizability tensor component $\alpha_{\rho\sigma}$ (ρ and σ designate the spatial coordinates x, y and z, as in equation 2.3) resulting from the molecular transition from state m to a virtual excited state and then to state n. An important aspect of equation 2.10 for the experimental Raman spectroscopist is that the scattered intensity is proportional both to the incident laser intensity, I_0, and to the fourth power of the frequency of the scattered Raman radiation, $\nu_0 - \nu_{mn}$. Since the incident laser frequency is adjustable and the frequency differences between vibrational states remain constant (a property of the molecule), the intensity of Raman scattering is essentially proportional to the fourth power of the incident laser frequency, ν_0. Therefore, a high-power, high-frequency laser will increase the intensity of Raman scattering.

The magnitude of the polarizability terms in equation 2.10 (*i.e.*, the set of $(\alpha_{\rho\sigma})_{mn}$ for all ρ, σ) determines the inherent Raman scattering intensity of a given transition. These polarizability terms are given by the Kramer Heisenberg Dirac (KHD) expres-

sion

$$(\alpha_{\rho\sigma})_{mn} = \frac{1}{h}\sum_{e}\left(\frac{\langle\Psi_n|\mu_\rho|\Psi_e\rangle\langle\Psi_e|\mu_\sigma|\Psi_m\rangle}{\nu_{em}-\nu_0-i\Gamma_e}+\frac{\langle\Psi_e|\mu_\rho|\Psi_m\rangle\langle\Psi_n|\mu_\sigma|\Psi_e\rangle}{\nu_{en}+\nu_0-i\Gamma_e}\right) \tag{2.11}$$

where e denotes a vibrational level of an excited electronic state (recall figure 2.1), ν_{em} and ν_{en} are the frequencies corresponding to the energy differences between the subscripted states, and $i\Gamma_e$ is a small term (relative to the other frequencies in the denominators of equation 2.11) that is related to the lifetime of the excited state and determines the natural width of Raman spectral lines [109]. The integrals in the numerators of equation 2.11 (*e.g.*, $\langle\Psi_n|\mu_\rho|\Psi_e\rangle$) are the electric transition moments between the subscripted states, where Ψ is the total wavefunction of the subscripted state and μ is the electric dipole moment operator along the coordinate specified by ρ, σ [109, 110]. These integrals represent the mixing of a ground electronic vibrational state (Ψ_m or Ψ_n) with an excited electronic vibrational state (Ψ_e). The first term in equation 2.11 corresponds to the excitation from state Ψ_m to state Ψ_e via the integral $\langle\Psi_e|\mu_\sigma|\Psi_m\rangle$, followed by the decay from state Ψ_e to state Ψ_n via the integral $\langle\Psi_n|\mu_\rho|\Psi_e\rangle$. Since the molecule may initially exist in the excited electronic state Ψ_e, the second term in equation 2.11 accounts for the reverse process, returning the molecule to the state Ψ_e. However, since the denominator of the second term will always be large relative to the denominator of the first term (since $i\Gamma_e$ is small relative to $\nu_{em,n} \pm \nu_0$), the first term in equation 2.11 dominates in Raman scattering [109]. To determine the total change in the polarizability tensor component for a given m to n transition ($(\alpha_{\rho\sigma})_{mn}$), both terms in equation 2.11 are summed over all excited vibrational states e to describe the distorted electron configuration (*i.e.*, the virtual state in figure 2.1) resulting from the interaction of the excitation source (frequency ν_0) and the molecule. The sum of the squares of all $(\alpha_{\rho\sigma})_{mn}$ over each component of the polarizability tensor (equation 2.3) finally determines the inherent Raman intensity (equation 2.10) of a given vibrational state transition.

Further analysis of the first term of the KHD expression allows for an explanation of two important results, each of which are only briefly discussed here as the full treatment is available elsewhere [109–111]. The first phenomenon relates to resonance Raman scattering, which can result in a factor of 10^3 to 10^5 increase in Raman scattering intensity I_{mn}. If the excitation frequency ν_0 is tuned to the transition frequency ν_{em} to the excited electronic state, the denominator of the first term in equation 2.11 approaches $i\Gamma_e$ (a very small term), resulting in a large increase in

$(\alpha_{\rho\sigma})_{mn}$ and subsequently I_{mn}. The second point provides a mathematical explanation of the quantum mechanical selection rules for normal Raman scattering. If the total wavefunctions Ψ are separated into electronic, vibrational and rotational component functions and the electronic component is expanded about the equilibrium nuclear position, equation 2.11 becomes an expression dependent on integrals, describing vibrational state transitions, that are only non-zero if there is only one quantum of energy difference between the initial and final vibrational states (*i.e.*, $m - n = \pm 1$) [109]. This result selects the transitions between vibrational states that will result in a change in the polarizability tensor component $\alpha_{\rho\sigma}$, and therefore yield Raman scattering.

Evaluating the KHD expression for a given vibrational transition of a polyatomic molecule can be computationally intensive, and requires knowledge of the wavefunctions of all of the vibrational states involved (*i.e.*, Ψ_m, Ψ_n, and all Ψ_e). However, for relatively simple molecules, analysis of molecular normal mode vibrations (discussed below) can provide useful insight into Raman scattering and the determination of the Raman activity of vibrational transitions.

2.2.3 Molecular vibrations and Raman activity

The vibration of a polyatomic molecule can be very complicated. However, any vibration can be broken down into an orthogonal set of normal mode vibrations that oscillate independent of each other, and usually with a unique frequency. A complicated vibration pattern can then be expressed as a superposition of the normal mode vibrations; it is these normal mode frequencies that are detected by a Raman spectroscopy instrument, provided that the normal mode vibration is Raman-active. The number of normal modes depends on both the number of atoms in the molecule and the geometry of the molecule.

Each normal mode is a unique vibrational degree of freedom (DOF) for the molecule. An N-atom molecule has $3N$ DOF of motion, since any atom can move in any of the three independent directions x, y, and z. However, $3N$ includes three translational DOF for the molecule as a single unit, and three rotational DOF about the principal axes running through the molecule's centre of mass. For a linear molecule, there is no rotation about the molecular axis, so there are only two rotational DOF. As such, the number of normal mode vibrations is equal to $3N - 6$ for an arbitrary molecule, and $3N - 5$ for a linear molecule. Carbon dioxide (CO_2), for example, is a

linear triatomic molecule, and therefore has $3 \times 3 - 5 = 4$ normal mode vibrations.

It is common practice in vibrational spectroscopy to assign descriptive names to the various types of normal mode vibrations. As an example, the normal mode vibrations of CO_2 consist of one symmetric stretching mode (figure 2.2a), two bending or deformation modes (figure 2.2b), and one antisymmetric stretching mode (figure 2.2c). The only difference between the two bending modes (figure 2.2b) is that their planes of oscillation are perpendicular to each other. As such the two modes have the same vibrational frequency and are called doubly degenerate modes.

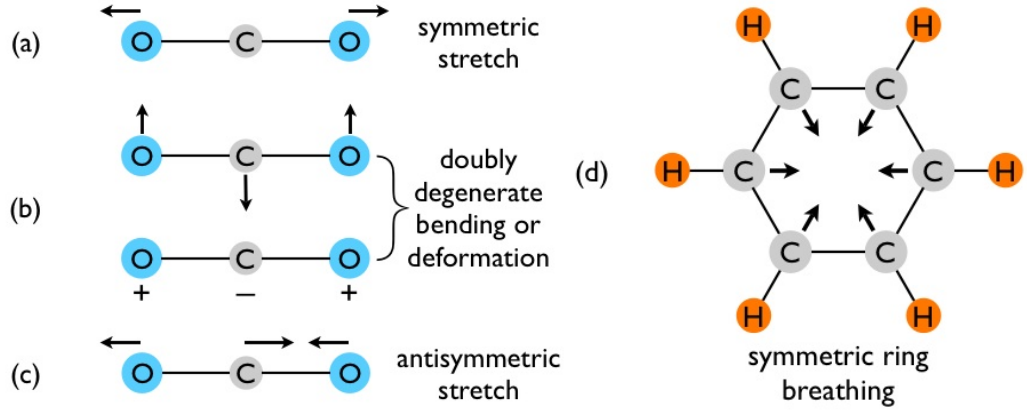


Figure 2.2: (a)-(c) Normal mode vibrations of the CO_2 molecule (O=C=O). In (b), + and − indicate motion in and out of the page, respectively. (d) Symmetric ring breathing of benzene (C_6H_6). Reproduced from Quinn Matthews' M.Sc. thesis [28].

For more complicated molecules (*i.e.*, $N > 3$), the normal modes become increasingly complex. However, a simple example of a large N molecule normal mode is the symmetric ring breathing mode of an aromatic ring, such as the one in benzene (C_6H_6) (figure 2.2d). Aromatic ring breathing modes typically yield intense Raman scattering and are often easily observed in chemical and biological Raman applications.

Not all normal mode vibrations produce Raman scattered photons. As discussed in section 2.2.1, a vibration is only Raman-active if there is a change in polarizability with a small nuclear displacement from the equilibrium position (*i.e.*, $(\delta\alpha/\delta q_i)_0 \neq 0$ in equation 2.8). As an example, consider the normal mode vibrations of CO_2. Due to the molecular geometry, the polarizability is not the same in all directions from the centre of mass of the molecule. For CO_2 the electrons are most polarizable (highest α) along the molecular axis, and are least polarizable perpendicular to the molecular axis. Plotting $1/\sqrt{\alpha}$ (by convention [110]) in all directions from the centre of mass of

the molecule creates a polarizability ellipsoid. If the size, shape or orientation of the ellipsoid is changing at the equilibrium position during the normal mode vibration, then the vibration is Raman-active. The changes of the polarizability ellipsoid for the three types of vibrations of CO_2 are depicted in figure 2.3.

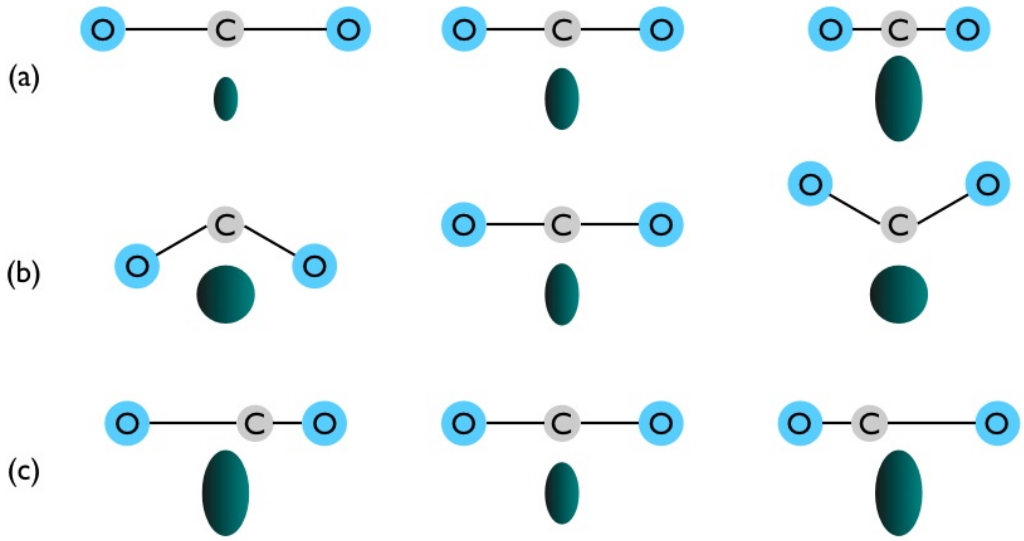


Figure 2.3: Changes in polarizability ellipsoids during normal mode vibrations of CO_2: (a) symmetric stretch is Raman-active, (b) bending or deformation is Raman-inactive, and (c) antisymmetric stretch is Raman-inactive. Reproduced from Quinn Matthews' M.Sc. thesis [28].

The symmetric stretching mode of CO_2 (figure 2.3a) is Raman-active, as the size of the polarizability ellipsoid is changing at the equilibrium position (getting larger or smaller); therefore $(\delta\alpha/\delta q_i)_0 \neq 0$. The bending or deformation mode (figure 2.3b) is Raman-inactive because the shape of the ellipsoid is the same on each side of the equilibrium position. Therefore, considering only small displacements at the equilibrium, $(\delta\alpha/\delta q_i)_0 = 0$. For the antisymmetric stretching mode (figure 2.3c), even though the size of the ellipsoid changes, the vibration is Raman-inactive because, again, the shape of the ellipsoid is the same on each side of the equilibrium position, and therefore $(\delta\alpha/\delta q_i)_0 = 0$.

Although useful for simple molecules, this method of determining Raman activity or inactivity by inspection of the normal modes is difficult to apply to large or complex molecules. However, one useful result of such analysis is that for molecules that have a centre of symmetry (*e.g.*, CO_2, benzene), symmetric normal mode vibrations are always Raman-active. Examples of such vibrations are the symmetric

stretching of CO_2 (figure 2.2a) and the symmetric ring breathing of benzene (figure 2.2d). Conversely, antisymmetric normal mode vibrations in such molecules are always Raman-inactive. Examples of such vibrations are the bending or deformation mode and the antisymmetric stretching mode of CO_2 (figures 2.2b and 2.2c). These conditions do not apply to molecules without a centre of symmetry (*e.g.*, H_2O), but in general, for any given molecule, symmetric normal mode vibrations are typically more intense in Raman scattering than antisymmetric vibrations [109].

2.3 Raman spectroscopy instrumentation

2.3.1 Raman shift

When measuring the frequencies of Raman scattered light, it is a universal convention to record the wavenumber, $\bar{v} = 1/\lambda$, rather than the frequency, v. Wavenumber is commonly expressed in units of cm^{-1}. Since in Raman spectroscopy it is the change in wavenumber from that of the excitation source that is of interest, the measured Raman signal is expressed as a "Raman shift", given by

$$\text{Raman shift} = \frac{1}{\lambda_0} - \frac{1}{\lambda} = \frac{v_0}{c} - \frac{v}{c} = \bar{v_0} - \bar{v} \quad (\text{cm}^{-1}) \tag{2.12}$$

where λ_0, v_0 and $\bar{v_0}$ correspond to the incident light and λ, v and $\bar{v}$ correspond to the Raman scattered light. Since $\Delta E = h\Delta v$, equation 2.12 shows that the Raman shift is proportional to the change in energy between the incident and Raman scattered light, and is thus proportional to the energy of the vibrational transition (recall figure 2.1). When creating a Raman spectrum, the Raman scattered light is dispersed by a spectrometer (discussed below, section 2.3.3) and captured on a CCD (discussed below, section 2.3.4), and the measured intensities are plotted against the Raman shift in wavenumbers.

2.3.2 Raman spectroscopy apparatus

The basic equipment and set-up of a typical RS apparatus is shown in figure 1.7a. A laser beam illuminates the sample and the Raman scattered light is collected and collimated by an objective lens. The Raman light is focused into a spectrometer, which disperses the light and focuses it onto a CCD detector. The general principles of light

dispersion in a spectrometer and light collection at a CCD detector are presented below in sections 2.3.3 and 2.3.4, respectively. General requirements and considerations for performing Raman microscopy are discussed in section 2.3.5. Details specific to the Raman system used in this work are presented in chapter 3.

2.3.3 Light dispersion

Once the Raman scattered light has been collected from the sample, the different scattered wavenumbers must be separated to record the full spectrum of information. In most modern Raman systems this is done using a dispersive spectrometer with a diffraction grating, with the exception being Fourier-transform Raman systems, which use infrared excitation (typically 1064 nm, rendering them impractical for Raman microscopy systems) and an interferometric approach to spectrum generation [112]. There are many different dispersive spectrometer designs in use, depending on the excitation wavelength, the type of RS being performed, and the spatial and spectral resolution requirements of the user. The dispersive spectrometer shown in figure 2.4 is a variation on a common design (single grating Czerny-Turner spectrometer [112]), and is the spectrometer design in the Raman microscope used in this work. The collected Raman scattered light, composed of a mixture of wavenumbers (*e.g.*, $\bar{\nu}_1$, $\bar{\nu}_2$, $\bar{\nu}_3$), is focused through the entrance slit of the spectrometer (recall figure 1.7a), passed through a collimating lens, and reflected by a plane mirror onto a diffraction grating. The diffraction grating reflects the light at different angles depending on the wavelength, and the dispersed light is reflected by another plane mirror through a focusing lens onto the exit focal plane. In this manner, each wavenumber contained in the Raman scattered light that passes through the entrance slit is focused to a different point along the length of the exit focal plane (figure 2.4).

The extent to which the light is spread across the exit focal plane of a spectrometer is described by the reciprocal linear dispersion, defined as the range of wavelengths or wavenumbers contained within a unit length of the focal plane. Reciprocal linear dispersion is commonly expressed in units of nm/mm if using wavelength, or cm^{-1}/mm if using wavenumber. In terms of wavelength, the reciprocal linear dispersion is given by

$$\frac{d\lambda}{dx} = \frac{10^6 \cdot \cos\theta}{n \cdot l \cdot F} \tag{2.13}$$

where n is the diffraction order, l is the line density of the grating in lines per mm (lines/mm), F is the exit focal length of the spectrometer in mm, and θ is the angle

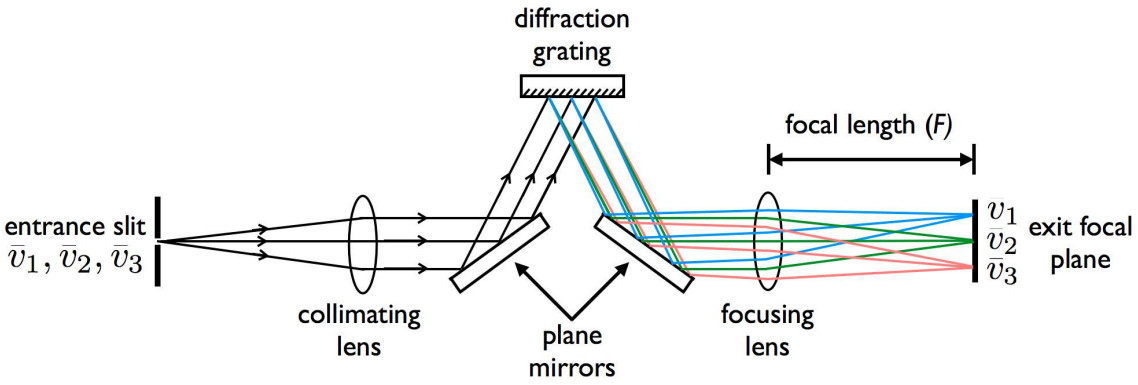


Figure 2.4: Design of the dispersive spectrometer used in this work. Light rays composed of a mixture of wavenumbers (*e.g.*, $\bar{v}_1$, $\bar{v}_2$, $\bar{v}_3$) are dispersed by a diffraction grating and focused onto the exit focal plane.

of the diffracted light leaving the grating [112, 113]. Equation 2.13 can be expressed in terms of wavenumber by using

$$\frac{d\lambda}{dx} = \frac{d(1/\bar{v})}{dx} = \frac{-1}{\bar{v}^2} \cdot \frac{d\bar{v}}{dx} \tag{2.14}$$

and therefore

$$\frac{d\bar{v}}{dx} = -\bar{v}^2 \cdot \frac{d\lambda}{dx} = -\frac{10^6 \cdot \cos\theta \cdot \bar{v}^2}{n \cdot l \cdot F} \tag{2.15}$$

Therefore, the reciprocal linear dispersion in cm^{-1}/mm varies with the square of the wavenumber of the scattered Raman light, which leads to a non-linear relationship between wavenumber and unit length at the focal plane. An instrument with a higher linear dispersion (or conversely a lower reciprocal linear dispersion $d\bar{v}/dx$) will spread a given wavenumber range over a greater length of the focal plane, and will therefore more easily resolve fine spectral details. Such an instrument will therefore have better spectral resolution than an instrument with low linear dispersion (high $d\bar{v}/dx$). Spectral resolution is defined as the minimum separation between two spectral features that can be clearly resolved, and modern Raman systems achieve spectral resolutions on the order of $\sim$2 cm^{-1} [112]. Methods for determining spectral resolution experimentally are described elsewhere [28].

In all Raman systems there is a practical limitation to the physical size of the focal plane (*e.g.*, width of the CCD detector, typically $\sim$1 to 3 cm). Therefore, there is a limited wavenumber range, or "spectral window", that can be measured in a single acquisition. The size of the spectral window can be increased at the

expense of the linear dispersion by decreasing either the grating line density l or the focal length F (equations 2.13 and 2.15), but this may be undesirable due to the decreased ability to resolve spectral details. A common technique used to avoid this limitation is to select for a desired spectral window by rotating the diffraction grating such that only the diffracted light from the desired window is incident on the focal plane of the spectrometer. This method requires the collection of multiple spectra at different grating angles to cover a larger wavenumber range, but does not sacrifice the spectral resolution of the system. Some modern Raman systems employ continuous scanning techniques, which eliminate the need for multiple acquisitions to cover a large wavenumber range without sacrificing spectral resolution [114]; one such method is Renishaw's SynchroScan technique (Renishaw Inc., Illinois, IL, USA), which is used in this work and is described in chapter 3. Modern Raman systems will span a spectral window anywhere from $\sim$100 cm^{-1} to over 4000 cm^{-1} wide, depending on the system configuration [112].

2.3.4 Light detection and Raman spectrum creation

The current method for detecting the dispersed light output from the spectrometer is to use a charge-coupled device (CCD). CCDs are two-dimensional (2D) optical arrays of photosensitive diodes, usually composed of a silicon-metal-oxide semiconductor [107, 108]. Small metal pads are deposited on each photosensitive "pixel" (array element) and kept at a positive potential; each is connected via a diode to a grid circuit that defines the CCD array dimensions. When photons interact with a photosensitive element, photoelectrons are produced and are attracted to the nearest metal pad. The number of photoelectrons collected at the pad is therefore proportional to both the intensity of the incoming light and the acquisition time (the amount of time the photosensitive element is exposed). After each acquisition, the charge collected (*i.e.*, the accumulation of photoelectrons) at each pad is read out pixel-by-pixel by sequentially shifting the charges from one row (or column) to the next as the charges from the row (or column) at the edge of the array are read out. The charge measured at each pixel is converted to an analog voltage, passed through an amplifier, and digitized by an analog-to-digital converter. The digital output allows a 2D image to be created where each pixel value is proportional to the light intensity incident on the corresponding CCD element.

The CCD detector array is placed at the focal plane of the spectrometer (figure 2.4)

such that the long axis of the CCD is parallel with the direction of linear dispersion. A typical arrangement is depicted in figure 2.5, using the three wavenumbers $\bar{\upsilon}_1$, $\bar{\upsilon}_2$ and $\bar{\upsilon}_3$ shown in figure 2.4. Typically, the height of the dispersed focal spot on the CCD, in the direction perpendicular to the dispersion axis, is greater than the size of a single pixel (*e.g.*, ∼0.02 mm). This is illustrated in figure 2.5, where each wavenumber, represented by a single distinct colour, is incident on six vertical pixels in each column. To generate a Raman spectrum of maximum signal intensity, the signal from the six pixels in each column are summated, or "binned", into a single measurement of intensity for that wavenumber. This selective binning practice serves to maximize the intensity of the Raman scattering at each wavenumber, while minimizing the CCD readout noise from pixels outside of the illuminated area.

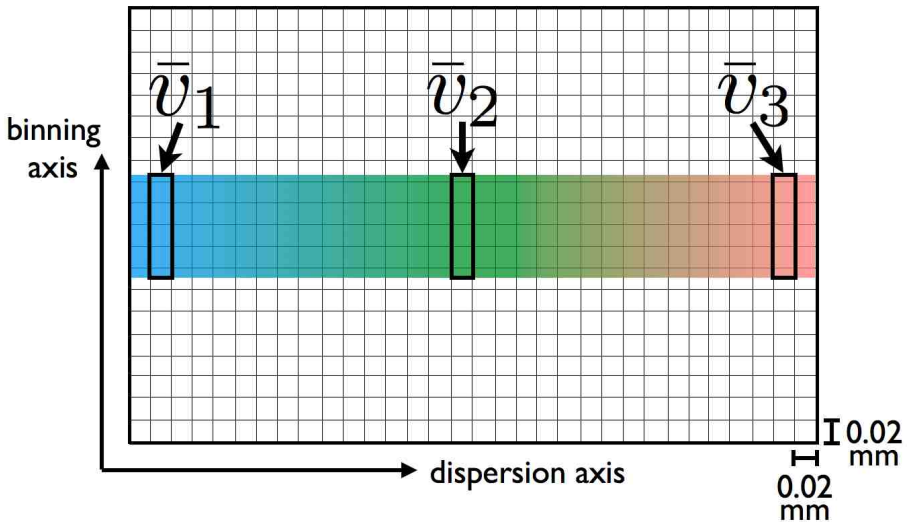


Figure 2.5: The arrangement of a CCD detector array placed at the focal plane of the spectrometer shown in figure 2.4, showing the alignment of the dispersion and binning axes with the CCD pixel grid. A typical pixel size is indicated.

There are two primary sources of CCD noise that contribute to the measured signal. Heat produces electrons in the photosensitive element that are indistinguishable from the photoelectrons produced by incident light, thus creating thermal noise, or "dark current". Thermal noise can be problematic for long acquisition times, but can generally be reduced to a negligible level by cooling the CCD detector, either thermoelectrically or with external coolants such as liquid nitrogen, to temperatures well below freezing (*e.g.*, −60 °C to −120 °C). Reading out the collected charges from the CCD also contributes noise, as a result of electrons produced while shifting charges from pixel-to-pixel and noise introduced during amplification and digitization. This

"readout noise" is independent of acquisition time, and for most Raman applications the signals detected are of much higher intensity than the readout noise level. Readout noise is generally much stronger than thermal noise for sufficiently cooled modern Raman systems, and can become problematic if the Raman signals are very weak. If necessary, post-processing techniques such as averaging and smoothing can be applied to reduce CCD noise.

The final Raman spectrum (*e.g.*, figure 1.7b) is generated by plotting the binned intensities (figure 2.5) against the known Raman shift corresponding to the position of each pixel along the dispersion axis of the CCD. Typically, the known Raman shifts across the CCD are calibrated by a sample with an accurately known Raman spectrum, such as tylenol or silicon [112].

2.3.5 Raman microscopy

Transforming a conventional Raman spectroscopy apparatus (figure 1.7a) into an effective Raman microscopy (RM) system requires the addition of a high degree of spatial localization of the Raman signal. In practice this is technically demanding, as the dimensions of the sample can be as small as a few microns across and less than a micron thick. The inherently weak signal in RS requires the use of a high-powered microscope objective to focus the incident laser beam to a very small spot in order to ensure efficient production and collection of the Raman signal, while not overheating the sample. In addition, to properly localize the point of Raman collection, a precisely controlled microscope stage with micron-scale stepping resolution is typically used.

Most RM systems couple the excitation laser into an upright microscope and collect the Raman scattered photons with the same objective that focuses the laser onto the sample in a 180° backscatter orientation (figure 2.6). This geometry is the most efficient method of Raman collection for thin, opaque samples, as it minimizes the loss of signal due to attenuation in the sample or substrate. It also allows the use of high-power objectives, which must be brought very close to the sample (*e.g.*, $0.2 - 0.4$ mm) for proper focus. In a 180° backscatter geometry, the Raman and Rayleigh backscattered light follow the same beam path as the incident laser. Optical filters remove the Rayleigh scattered light and pass the Raman scattered light to the spectrometer (figure 2.6).

Another important aspect of a RM system is confocal discrimination. The term confocal refers to the principle axis of an imaging system, which for RM is the axis

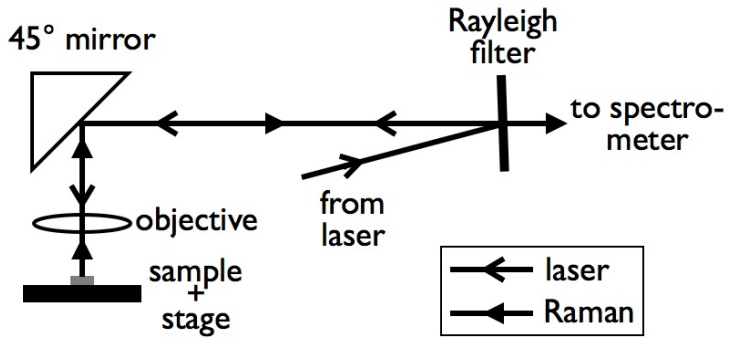


Figure 2.6: General schematic of the light path for a typical RM system using a 180° backscatter collection geometry.

along which both the incident laser and the collected signal traverse (see figure 2.6). Confocal microscopy involves placing a small spatial aperture (such as a pinhole or series of slits) at the focus of any conjugate image plane on the principle axis, after the signal has been collected but before the signal is recorded (figure 2.7). In RM systems the confocal aperture is usually placed after the Rayleigh filter and before the spectrometer. The confocal aperture serves to reject signals originating from an out-of-focus or off-axis region of the sample or RM system. Such unwanted signals could be stray light, scattered signals from outside the focal volume, or signals from a sample substrate. A smaller confocal aperture increases confocal resolution (*i.e.*, the depth selectivity in the sample) and reduces unwanted signals, but if the aperture is too small it may decrease the amount of Raman signal collected from the desired target. RM systems typically employ confocal apertures ranging from 10-200 μm in diameter.

The choice of microscope objective is essential for RM. The glass components in the objective must be of high optical quality to minimize fluorescence. In most RM systems the objectives used are infinity-corrected, so that a perfectly collimated light source (*i.e.*, an ideal laser, focused at infinity) incident on the back aperture of the objective is focused to a diffraction limited spot. The dimensions of the focused laser spot greatly affect the collection efficiency and spatial and confocal resolutions of the system, and depend on the characteristics of both the incident laser and the objective. The most important specification of an objective for RM is the numerical aperture (NA), which is defined as $\mathrm{NA} = n \cdot \sin\alpha$, where n is the refractive index of the immersion medium, and α is the half-angle of the maximum light cone collected by the objective (figure 2.8). Since $\alpha_{max} = 90°$, NA is always < 1 for dry objectives

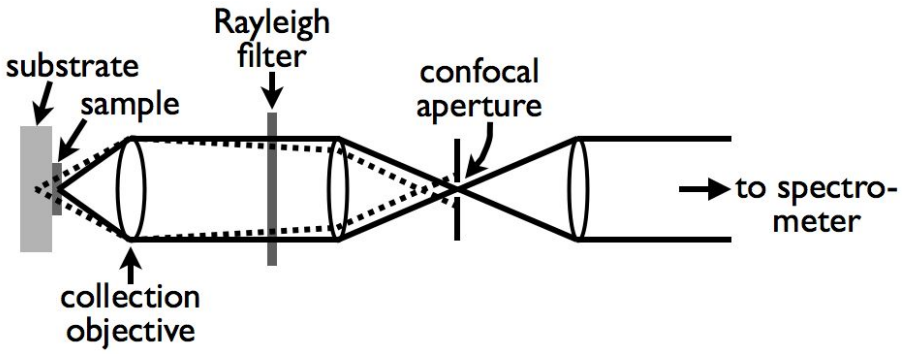


Figure 2.7: Principle of confocal Raman microscopy. Signals originating from an out-of-focus or off-axis position such as a substrate (dotted rays) are rejected by the confocal aperture, whereas signals from the desired sampling volume (solid rays) pass through the aperture to the spectrometer.

(no immersion, $n = 1.00$).

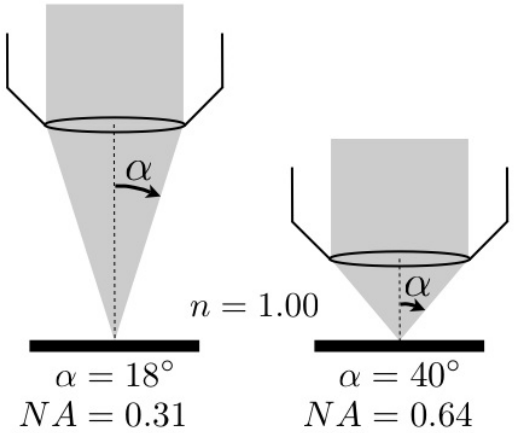


Figure 2.8: Example of numerical apertures (NA) calculated for two microscope objectives, using $\mathrm{NA} = n \cdot \sin\alpha$, in air ($n = 1.00$). Reproduced from Quinn Matthews' M.Sc. thesis [28].

The NA of an objective determines its collection efficiency, given by the solid angle intercepted by the objective $\Omega = 2\pi(1-\cos\alpha)$. Since the maximum value for α is 90°, the maximum solid angle Ω is 2π, resulting in a maximum efficiency of 50% ($2\pi/4\pi$, where 4π is the solid angle subtended by a complete sphere). A higher NA objective is more efficient at collecting Raman scattering but also excites a smaller volume of the sample (shown below). Therefore, optimizing signal strength depends on both the properties of the objective and the dimensions of the sample.

The minimum width of the focused laser spot depends on the NA of the focusing

objective and the laser wavelength, and can be approximated using the formula [115]

$$s = \frac{0.61 \cdot \lambda}{\mathrm{NA}} \tag{2.16}$$

The height of the laser spot along the beam axis is often called the depth-of-focus (DOF), and can be approximated by [115]

$$DOF = \frac{n \cdot \lambda}{\mathrm{NA}^2} \tag{2.17}$$

Equations 2.16 and 2.17 are working formulas derived using wave optics [115], and are only valid for perfectly collimated lasers that are perfectly aligned with the back aperture of a high NA objective. This scenario is never achieved in practice due to laser beam divergence ($\sim$4 milliradians for the 785 nm laser used in this work, but highly dependent on laser model and wavelength) or optical misalignments, but the formulas are useful to demonstrate how the choice of microscope objective affects the size of the Raman excitation and collection volume. Maximum spatial and confocal resolution is achieved with the smallest possible focused spot, and therefore is achieved with a high NA objective and a short wavelength laser, aided by a small confocal aperture (recall figure 2.7). RM systems using high NA (0.7-1.3) objectives and laser wavelengths in the $500 - 650$ nm range can achieve a spatial resolution as low as 0.5 μm, and a confocal resolution as low as 2-3 μm.

2.4 Raman spectroscopy of cells

There are several unique considerations to address when performing RS on cells or tissues. The most important considerations are (1) the choice of laser wavelength and power, (2) the spatial and confocal resolution requirements, (3) the spectral resolution requirements and spectral window(s) of interest, and (4) the choice of substrate material. The specific materials and methods used to address these considerations in this work are presented in chapter 3.

2.4.1 Laser wavelength and laser power

The choice of laser wavelength and power is vital when performing RS analysis of cells. A higher laser power yields a higher laser intensity at the sample, and the intensity of Raman scattering is proportional to both the intensity of the laser and the fourth

power of the laser frequency (recall equation 2.10). Therefore, a high power, short wavelength laser generates the strongest Raman scattering. However, two factors that mediate the choice of wavelength and power are biological fluorescence and sample photodamage.

Biological fluorescence can mask weak Raman signals, but is significantly reduced when using long visible or near-infrared wavelengths (*i.e.*, $\geq$ 650 nm) [110, 116]. The least amount of fluorescence induced by laser light in the visible spectrum occurs for near-infrared wavelengths (*i.e.*, $>$ 700 nm). The reduction in fluorescence at these wavelengths comes with an undesirable reduction in Raman scattering intensity compared to shorter wavelengths (*i.e.*, 450 - 650 nm), but allows for longer acquisition times and the detection of weaker Raman signals. This effect was reported by Schuster *et. al.* [65] during analysis of single bacteria cells, where Raman signals collected using 514 nm excitation were dominated by fluorescence, yet at 632.8 nm excitation the Raman signals were clearly detected above a much reduced fluorescent background signal. 783 nm excitation produced the least fluorescence, but no biological Raman signals were observed due to insufficient laser power density and signal collection efficiency [65].

Laser-induced photodamage can be determined experimentally by observing a changing Raman signal as a function of illumination time. Several studies have shown that for short visible wavelengths such as 488 nm and 514 nm, photodamage to cells and chromosomes is observed after only a few minutes for laser powers at the sample of $<$ 10 mW [57, 58, 117]. It should be noted that the power density (*e.g.*, in units of mW/μm^3) at the sample is really the limiting factor (as opposed to power in mW), and local damage to the sample from heat and 2-photon processes may be caused if the power density is too high; however, power density is often not reported in the literature. The photodamage observed with $<$ 10 mW of 488 nm and 514 nm light was attributed to photochemical reactions, as the laser powers used were deemed much too low to induce any sample heating [57]. The same effect is not observed for 660 nm, where cells were reported to survive for up to 20 minutes at 25 mW laser power. Furthermore, it has been shown that even longer wavelengths allow the use of much higher powers and longer acquisition times, as cells examined at 785 nm exhibit no spectral changes even after 40 minutes at 115 mW laser power [58]. Cells have also been analyzed at 830 nm at over 100 mW laser power for 20 seconds [79]. These results suggest that the best range of laser wavelengths for single cell RS applications is 785 - 830 nm, which is also a desirable wavelength range for minimizing the effect

of biological fluorescence.

2.4.2 Spatial and confocal resolution

The spatial and confocal resolution of a RS system should be chosen carefully depending on the application. For investigating the interior structures of a single cell adhered to a growth substrate (where cells are typically 10-50 μm in diameter and 1-5 μm thick), a spatial resolution on the order of 1 μm is desirable. In this case, the confocal resolution should be on the order of the cell's thickness to maximize the efficiency of signal collection at the point of acquisition, as adherent cells are usually "splayed out" on the substrate and therefore have very little structural variation with depth at a given point. However, for acquiring a single representative Raman spectrum from an entire cell, either from a suspension or from a pellet where cells assume a roughly spherical shape (diameter $\sim$10 μm), the spatial and confocal resolution should be on the order of the dimensions of the cell.

2.4.3 Spectral windows and spectral resolution

There are two primary spectral windows of interest for RS of cells: a low wavenumber window, $\sim$600 - 1800 cm^{-1}, which contains abundant spectral information from proteins, lipids, nucleic acids and carbohydrates [59, 65, 66, 75, 79, 81, 92, 118–121], and a high wavenumber window, $\sim$2700 - 3100 cm^{-1}, dominated primarily by contributions from CH stretching in proteins and lipids [75, 92, 119–121]. To date, the smallest spectral separation, as reported in the literature, between two distinct Raman peaks in either of these spectral windows is 6 cm^{-1} [59]. Therefore, to facilitate the detection of as many Raman features as possible, the spectral resolution of the RS system used for cellular analysis should be < 6 cm^{-1}.

2.4.4 Substrate material

The substrate material chosen should be transparent to the incident laser light in order to minimize the contribution from reflected or scattered light originating from outside the desired focal volume. A further consideration is that the substrate itself must be relatively Raman inert in the two primary spectral windows of interest for biological RS, discussed above. Two common sources of substrate background signal are unwanted Raman signals, and fluorescence caused by the incident laser exciting

the substrate material. Background signal from substrates is a recurring problem in Raman spectroscopy, and can be corrected for by various baseline subtraction methods [122] as long as the background features are broad and reproducible. However, as biological Raman signals are quite weak, a very low background signal from the substrate is nonetheless desirable. Several different substrates have been investigated for RS analysis of single cells, such as calcium fluoride [65, 120], barium fluoride [123], and low-fluorescence quartz [65, 119].

2.5 Advantages and disadvantages of Raman spectroscopy

2.5.1 General advantages

Applications of RS are found in a wide range of disciplines, some of which are mentioned below in section 2.6. The main strength of RS is the ability to obtain measurements of molecular composition, structure and dynamics, while working with optical wavelengths. Virtually any type of sample can be analyzed (provided the induced molecular vibrations are Raman-active), whether it be solid, liquid, gas or vapour. Aqueous molecules and biological specimens requiring water immersion can be analyzed with RS due to the very weak Raman signals from water. The presence of water is a major drawback to infrared (IR) absorption spectroscopy due to the high opacity of water in the IR spectrum. RS is also noninvasive and nondestructive for most samples, provided that the laser wavelength and power are chosen with care, the sample illumination time is properly considered, and undue sample heating is avoided. Most RS applications require minimal sample preparation. In addition, modern optics and lasers incorporated into a Raman microscopy system (figure 2.6) allow for spatial and confocal resolutions on the sub-micron to micron scale. Finally, Raman vibrations typically yield sharp, discrete signals, inherently providing a high level of spectral resolution compared to fluorescence-based analysis methods [110].

2.5.2 General disadvantages

The primary limitations of RS arise from the inherently weak Raman effect, as compared to, for example, fluorescence. As such, many Raman experiments require long integration times to achieve a sufficient signal-to-noise ratio (SNR). The detection

of Raman signals often competes with fluorescence or unwanted Raman signals from substrates, sample containers, optical components in the Raman system, or other substances or impurities in the sample itself. Fluorescence spectra are broad and continuous, and contain little information about the vibrational state of a molecule. Due to the inherently higher intensity, fluorescence signals can swamp Raman signals if the respective spectral windows overlap.

2.5.3 General considerations for biological samples

RS has unique advantages and disadvantages for investigating biological samples. The advantages of RS for biomedical applications were discussed in section 1.3.1, and these aspects also apply to many other biological samples. One disadvantage is that biological RS signals rarely originate from a single type of molecule, resulting in poor inherent molecular specificity as compared to other established biological analysis techniques (further discussed in section 2.5.4 below). Other disadvantages of RS for biological analysis are the inherently weak intensity of biological Raman scattering, often necessitating long integration times, and the competition with fluorescence from biological materials, optical components, and sample substrates or containers. However, the use of high-power excitation lasers, low-fluorescence optics and substrates, confocal microscopy systems and baseline removal methods [122] can alleviate these difficulties.

One important aspect of biological RS is that the spectra obtained are often very complicated. Spectra from complex biological samples (*i.e.*, cells or tissues) contain many overlapping spectral features that originate from different classes of biomolecules (*e.g.*, figure 1.7b). As such, it can be difficult to interpret the biological nature of changes in spectra from a given sample, or the biological reason for any differences in spectra between samples. However, this very difficulty arises because there are spectral features from many different biomolecules, which potentially allows for the simultaneous analysis of many different biomolecule types at once. Multivariate analysis techniques, such as principal component analysis (PCA), linear discriminant analysis (LDA) or cluster analysis, can aid in the interpretation of large biological RS data sets [68–80, 84]. Such techniques can potentially capitalize on this spectral complexity, and can significantly aid in the interpretation of biological RS results.

2.5.4 Comparisons with established molecular analysis techniques for cells or tissues

RS has specific advantages and disadvantages as compared to established molecular analysis techniques for cells or tissues, such as flow cytometry, positron emission tomography (PET) and magnetic resonance spectroscopy (MRS). A brief discussion of each of these techniques, with respect to the capability for molecular analysis of cells or tissues, relative to RS, is provided below. A general comparison of the attributes of each technique is included at the end of this section in table 2.1.

Flow cytometry

Flow cytometry, described in section 1.2.1 with respect to cell cycle analysis, involves treating cells in suspension with a fluorescent label that binds to a biomolecule of interest. The cell suspension is passed through a single-cell wide capillary and illuminated with an excitation laser, and a photo-multiplier tube records the fluorescent intensity from each cell, which is proportional to the content of the labelled molecule in a cell. These labels can be specific to a certain type of molecule (*e.g.*, just DNA or RNA [18]) or specific to a broad class of molecules (*e.g.*, all proteins [17] or all lipids [60]). Using a flow cytometer with multiple excitation lasers (up to 4 in modern systems), cells can be treated with multiple labels for the simultaneous analysis of relative amounts of different molecules, provided that the fluorescent emission spectra of the multiple labels do not overlap.

The primary advantage of flow cytometry compared to RS is that thousands of cells can be analyzed per second, allowing for very fast screening of a large population of cells. Furthermore, fluorescent labels can be very specific to a certain molecule, allowing for molecular specificity and sensitivity at small concentrations unachievable with standard RS techniques. However, flow cytometry is limited by the number of fluorescent labels that can be used at once (*i.e.*, the number of excitation lasers available), and by the ability to simultaneously resolve several distinct fluorescence emission spectra [17]. Flow cytometry also requires a suspension of cells, and therefore cannot be applied *in vivo*, or applied directly to excised (*ex vivo*) tissue samples without further extraction of a single cell suspension. Often flow cytometry requires pre-treating the cells in order to permeate the cell membrane (thus killing the cell) to allow for the fluorescent label to bind to intracellular biomolecules (*e.g.*, DNA). Furthermore, flow cytometry does not directly provide information about molecular

structure, which is an important strength of RS for biological analysis.

Positron emission tomography

Positron emission tomography (PET) is a molecular imaging modality in which a patient is injected with a positron-emitting label designed to bind to a specific molecule that circulates though a patient's vasculature. If the target molecule preferentially accumulates in specific organs or disease sites (*e.g.*, cancer), the detection and localization of the emitted positrons enables PET to be a powerful *in vivo* diagnostic and functional imaging modality [124]. However, for molecular analysis purposes, PET can only be used to detect one type of molecule at once, unless multiple positron-emitting labelled molecules are known to specifically accumulate in different parts of the patient anatomy that can be spatially distinguished. PET is also limited, like flow cytometry, by the requirement to attach a label to the molecule of interest. There are a limited number of molecules that can be labelled and specifically accumulated by anatomic sites of interest, although many such label-molecule conjugates have been recently developed for PET analysis of amino acid metabolism [125] and gene expression products [126]. However, like flow cytometry, these labels are often very specific to a certain molecule, providing molecular specificity and sensitivity at small concentrations unachievable with standard RS techniques. PET has the further advantage over RS of inherently being an *in vivo* technique. However, PET is not suitable for molecular analysis of single cells or sub-millimeter scale tissue sections as the spatial resolution limit of PET is on the order of 1 mm [127], as compared to 1 μm for Raman microscopy systems. In addition, PET cannot directly provide information about molecular structure.

Magnetic resonance spectroscopy

Magnetic resonance spectroscopy (MRS) utilizes differences in the magnetic properties of chemical nuclei in different molecules to obtain a spectrum where each peak corresponds to a nuclear resonance frequency shift specific to a certain molecule. Different nuclei are targeted for analysis of different molecules; for example, ^{1}H-MRS allows detection of metabolites such as citrate, creatine, choline, lactate, alanine, glutamine, myo-inositol and N-acetylaspartate [128–131], ^{31}P-MRS allows specific detection of phospholipid molecules such as phosphatidylcholine, phosphatidylethanolamine and phosphatidylinositol [60], and ^{13}C-MRS is often used to isolate glucose [124]. Like

PET, and unlike RS, MRS is easily applied *in vivo* using conventional MRI scanning technology, although the relatively low magnetic field strengths in clinical MRI units (typically 1.5-3.0 T), as compared to experimental small-animal MRI units (up to 21 T), currently limits the sensitivity and spatial resolution of the technique [128]. Like RS, MRS allows the simultaneous analysis of multiple types of molecules, but MRS is inherently limited to molecule types that are relatively abundant and yield a distinct MRS signal, such as those mentioned above. MRS can provide structural information, but this often requires the analysis of multiple nuclear targets (*e.g.*, ^{1}H, ^{13}C and ^{15}N for determination of protein and nucleic acid structure [132]), therefore requiring multiple acquisitions and complicated analysis methods (relative to the analysis of RS spectra). Like RS, MRS can be easily applied to *in vitro* and *ex vivo* samples, using smaller laboratory MRS units. However, unlike RS, MRS cannot provide the spatial resolution or sensitivity required for single cell analysis, as samples containing on the order of 10^6 cells are typically required for sufficient signal production [133].

Table 2.1: Comparison of Raman spectroscopy (RS) with flow cytometry, positron emission tomography (PET) and magnetic resonance spectroscopy (MRS), for the molecular analysis of cells or tissues.

	RS	**Flow Cyt.**	**PET**	**MRS**
Spatial resolution limit	∼.001 mm	one cell	∼1 mm	∼0.1 mm
Acquisition time	10 - 60 s	∼.001 s	∼1 s	60 - 1800 s
Detection sensitivity	Low - Moderate	High	High	Low
Molecular specificity	Low	High	High	Moderate
Molecular structure	Yes	No	No	Yes
In vivo applicability	Low	None	High	Moderate
Ex vivo tissues	Yes	No	No	Yes
Ex vivo cell suspensions	Yes	Yes	No	Yes
Single cells	Yes	Yes	No	No
Multiple molecule analysis	Yes	Limited	Limited	Yes

2.6 Example applications of Raman spectroscopy

The uses of RS are too numerous to fully discuss here, and are addressed in detail in various texts [109, 110, 112]. However, several major fields of application are mentioned here by way of example.

RS has been used extensively to analyze the structural chemistry of inorganic compounds, minerals and organic molecules. A few examples include determining the structural forms of amorphous carbon and diamond [134], distinguishing between

the A, B, and Z structural forms of double-stranded DNA [135], and identifying trace elements and minerals in urban dust [136]. Development of fiber-optic and remote Raman collection techniques have enabled the nondestructive chemical analysis of paintings, ancient dyes, and archeological artifacts [137]. A few major industrial applications of RS include development and characterization of polymers and emulsions [138], quality control of the mechanical and structural properties of semiconductors used in electronic devices [139], and forensic identification of illicit drugs and explosives [140]. In addition, RS is routinely applied in the pharmaceutical industry for drug screening and investigation of molecular polymorphism [141]. RS has also been used by the commercial food industry to analyze trace component levels, fatty acid content, molecular structures, and changes in food quality [142].

The use of RS in the biological sciences has become incredibly widespread due to the advantages discussed in sections 1.3.1 and 2.5. A thorough review of the many biological applications of RS can be found elsewhere [143, 144]. However, a few examples of the various biological RS applications over the past few decades are mentioned here.

Early Raman microscopy techniques were successfully used to measure the resonance Raman spectra of visual pigments, and their photoproducts, from single rod and cone photoreceptor cells in goldfish [145], enabling the analysis of pigment content that was previously impossible without difficult extraction of cells from the retina. Further demonstration of the noninvasive and nondestructive potential of RS was shown by a pair of Fourier-transform RS studies; one study performed analysis of the structure of water, proteins and lipids in fully intact samples of human hair, nails and skin [146], and another study performed measurements of organic and inorganic compounds in samples of ancient (4^{th} - 16^{th} century) and modern human teeth [147]. RS has also found utility in botanical applications, such as to determine the amount and types of naturally occurring carotenoids in a variety of edible plants [148]. A recent novel biomedical RS study performed noninvasive tomographic reconstruction of bone tissue within an intact section of a canine limb, via detection of RS signals from bone acquired through layers of skin and subcutaneous tissues [149].

There have been many modern applications of both RS and RM in single cell analysis and in radiobiological investigations, and these studies are directly relevant to this work. As such, these studies were discussed in sections 1.3.2 and 1.3.3, respectively.

Chapter 3

Materials & Methods

This chapter describes the experimental equipment and procedures used throughout the studies presented in the results and discussion chapters 4 to 7. Materials and methods specific to a certain chapter are presented therein. The first part of this chapter (section 3.1) presents the types of cells used in this work, and the general protocols for cell culture, sample preparation for Raman spectroscopy, cell irradiation and flow cytometry. The second part of the chapter (section 3.2) describes the design and operation of the Raman system used for this work, and the protocols for acquisition of single-cell Raman spectra. The final part of the chapter (section 3.3) presents the methods used for processing single-cell Raman spectra and for performing principal component analysis (PCA). Descriptions of the relevant theories and algorithms are provided, along with some examples of implementation.

3.1 Cells and cell processing

3.1.1 Human tumour cells

Seven human tumour cell lines were used in this work, derived from epithelial prostate (DU145, PC3, LNCaP and PacMet), breast (MDA-MB-231 and MCF7) and lung (H460). These cell lines are all adherent, meaning they attach to a surface for growth, but vary in growth morphology, as seen in the optical images of the cell lines (except PacMet) shown in figure 3.1, and by culture doubling time (table 3.1). These cell lines also vary by p53 gene status and intrinsic radiosensitivity, which is further discussed in chapter 7. Original cell stocks were obtained from American Type Culture Collection (ATCC, Manassas, VA, USA), with the exception of the PacMet cells (originally

characterized in 2008 by Troyer *et. al.* [150]), which were obtained from (and used with permission by) Dr. Linda deGraffenried at the University of Texas Health Science Center, Department of Medicine.

Table 3.1: Doubling time ($\mathbf{T}_d$) for the seven human tumour cell lines used in this work (data from American Type Culture Collection (ATCC, Manassas, VA, USA) unless otherwise indicated).

Cell Line	$\mathbf{T}_d$ (hrs)
H460	23
MCF7	29
MDA-MB-231	39 [151]
PC3	27 [152]
DU145	34 [153]
LNCaP	34
PacMet	24 [150]

3.1.2 Cell culture

To maintain a sterile environment, all cell culture work was done in laminar flow biological safety cabinets at the Deeley Research Centre (DRC) in the BC Cancer Agency Vancouver Island Centre (BCCA-VIC). The cells were cultured in 10-cm diameter tissue culture dishes in 15 mL of growth media, supplemented with 10% fetal bovine serum (FBS) (HyClone, Logan, UT). Dulbecco's Modified Eagle Medium (DMEM) media (HyClone) was used for the DU145, PC3, MCF7 and MDA-MB-231 cell lines, and Roswell Park Memorial Institute (RPMI) 1640 media (HyClone) was used for the H460, LNCaP and PacMet cell lines. To promote growth, cells were kept in an incubator at 5% CO_2 and 37 °C.

Every 3-4 days, or when the cells grew to 70-80% confluency within the dish, thus restricting their growth, the cell cultures were harvested and diluted into fresh media to allow the cells to continue healthy growth. The cells were harvested by first discarding the old media and rinsing the cells in ∼10 mL of phosphate buffered saline (PBS) (HyClone), an ionic buffer solution that maintains osmotic pressure within the cells. About 2 mL of trypsin (HyClone), an enzyme that causes cells to detach from the growth substrate, was added to the dish and the cells were incubated for ∼5 minutes, or until the cells were observed to fully detach from the dish surface. The trypsin was de-activated with FBS by adding 8 mL of fresh growth media, and the cell solution was centrifuged at 1200 revolutions per minute (rpm) for 5 minutes. The

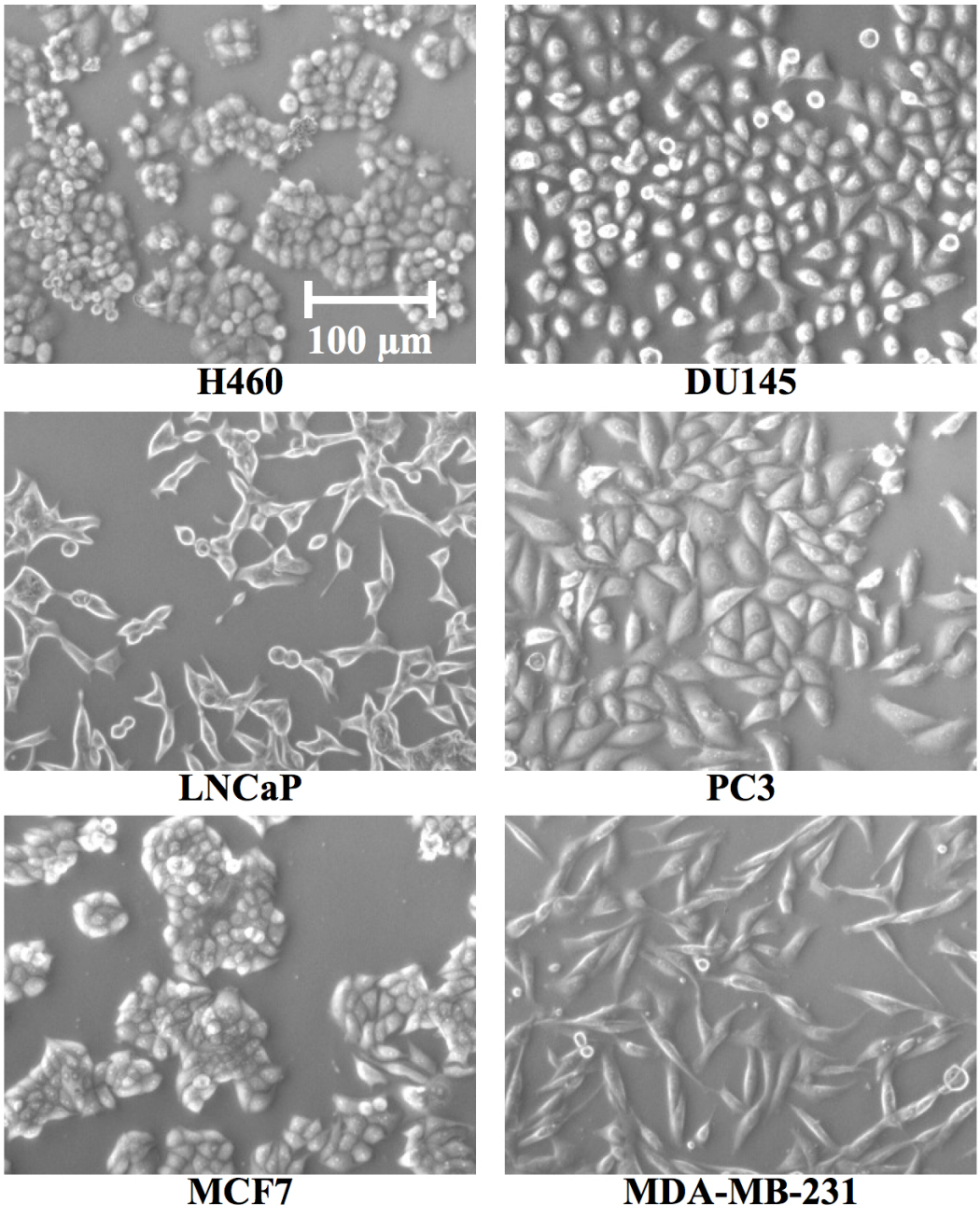


Figure 3.1: Optical images of H460, DU145, LNCaP, PC3, MCF7 and MDA-MB-231 human tumour cells adhered to the cell culture dish, obtained with a 20x objective on an inverted microscope. Field of view is $\sim 385 \times 290$ μm for all images.

supernatant was discarded and the cells were resuspended in 1 mL of fresh growth media. Depending on the level of dilution desired, a fraction of the cell suspension (*e.g.*, 100 μL for a 1:10 dilution) was transferred to a new dish with 15 mL of fresh culture media. The new cell suspensions were pipetted several times to ensure an even distribution of cells throughout the dish, and the dish was returned to the incubator.

3.1.3 Sample preparation for Raman spectroscopy

To prepare cells for Raman analysis, cultures were harvested by rinsing with PBS to remove dead cells and debris, adding trypsin to detach the remaining live cells, adding growth media with FBS to de-activate the trypsin, and centrifuging to discard the trypsin-media supernatant. Cells were then re-suspended in growth media and centrifuged into a pellet in a 200 μL plastic vial. Vials were kept on ice until RS analysis (1-6 hours), upon which the chosen pellet was transferred to a low-fluorescence quartz disc (Technical Glass Products, Painesville, OH, USA) in order to minimize spectral contributions from the sample substrate. The protocol for single-cell RS acquisition from the cell pellet is provided in section 3.2.2 below.

3.1.4 Cell irradiation experiments

For the primary cell irradiation experiments presented in chapters 6 and 7, 12 identical cell cultures were prepared in T-75 tissue culture flasks (75 cm^2 surface area for cell growth, physical dimensions $\sim$8.7 x 13.1 x 3.8 cm^3) at an initial confluency of 10-15% after sub-culturing. Prior to irradiation, cells were allowed to incubate for $\sim$3.5 days, growing to a confluency of 70-80%. Cell cultures were irradiated in T-75 flasks with a single fraction of 6 MV photons from a Varian 6EX linear accelerator (Varian Medical Systems, Inc., Palo Alto, CA, USA), at a dose rate of $\sim$5.9 Gy/min. Five cultures were irradiated to 50 Gy, one was irradiated to 15 Gy, one was irradiated to 30 Gy, and the remaining five cultures were left unirradiated. All flasks were removed from the incubator at the same time and were returned to the incubator within 30 minutes of removal. Cultures irradiated to 50 Gy were harvested for RS analysis at 0, 24, 48, 72 and 120 hours post-irradiation, with one unirradiated culture harvested and analyzed at the same time (figure 3.2). Cultures irradiated to 15 and 30 Gy were harvested and analyzed with RS at 72 hours post-irradiation (figure 3.2). The 120 hours post-irradiation cultures were only analyzed with RS if the viability of both cultures (measured with flow cytometry, see section 3.1.5 below) exceeded

75%. Similar protocols were followed for the other irradiation experiments presented in chapter 6, with the relevant differences provided therein.

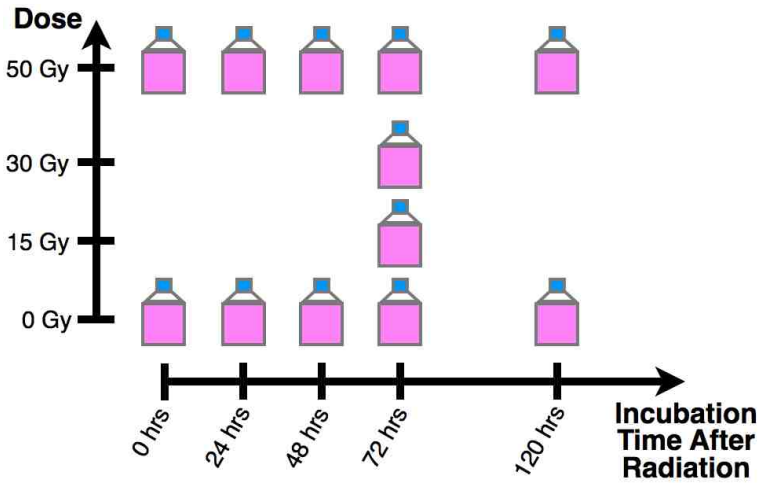


Figure 3.2: Protocol for irradiated dose and time of RS analysis post-irradiation, for cell culture irradiation experiments using T-75 tissue culture flasks.

The linac set-up for the cell culture irradiations is shown in figure 3.3. Six flasks were placed on top of a 5 cm thick 40 x 40 cm^2 block of water equivalent plastic (Standard Grade Solid Water®, Gammex, Inc., Middleton, WI, USA) in a 2 x 3 flask arrangement centred on the photon beam axis, with the bottom of all flasks resting on top of the block. The linac gantry angle was set at 180° in order to irradiate the cells from the bottom, through the 5 cm block, to ensure sufficient build-up of dose deposition prior to the flasks [10]. This arrangement also minimized any extra air gap between the radiation source and the cells, which were adhered to the bottom of the flasks and covered by a thin layer (∼3 mm) of growth media. An identical water equivalent plastic block was placed on top of the flasks as well. The top surface of the bottom block was set at 100 cm from the beam source, such that the photon beam axis intersected with the plane formed by the bottom of the flasks at the linac isocentre (recall chapter 1, figure 1.1). The linac jaws were set to deliver a field size of 30 x 30 cm^2 at the isocentre, to provide as uniform dose coverage as possible over all six flasks (∼26 x 26 cm^2 total area). For experiments requiring the flasks to be irradiated to varying doses, a given flask was replaced with an empty flask (or other flask containing cells to be irradiated) after the desired dose for the flask was delivered. Dose calculations for this set-up were provided by Dr. Will Ansbacher, using the Vancouver Island Cancer Centre's "MU Calculator", version 4.74.

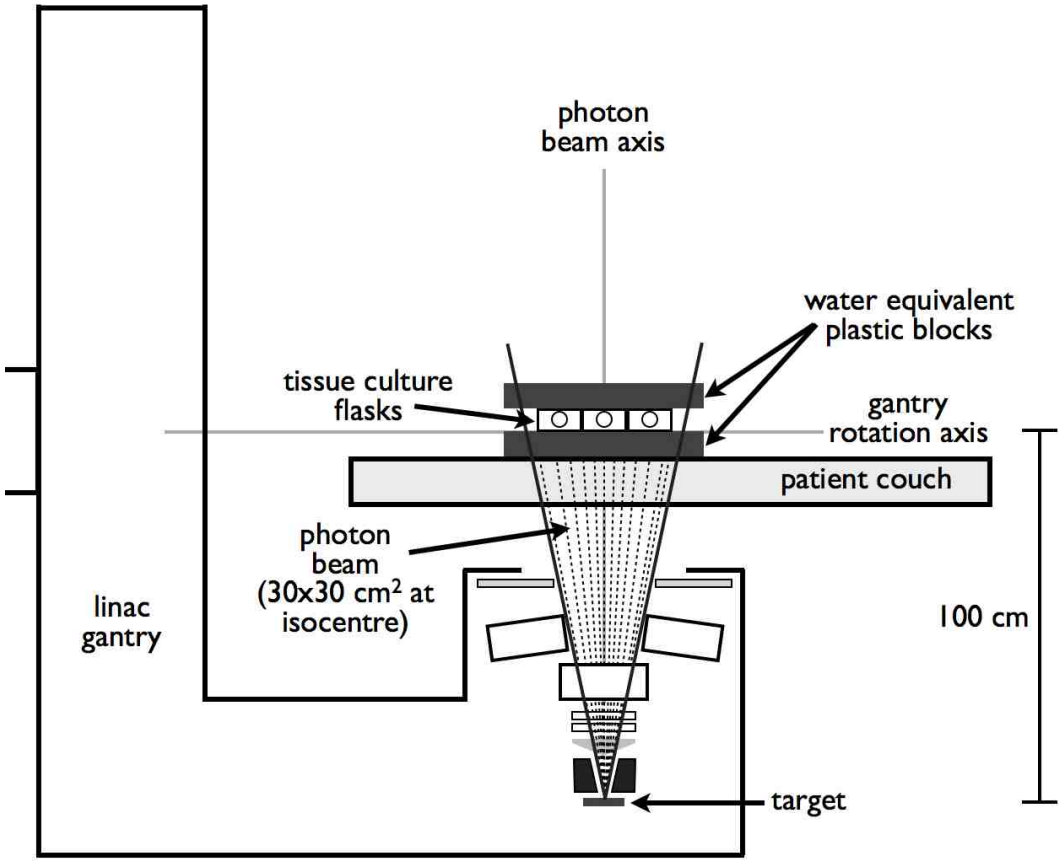


Figure 3.3: Linac set-up for cell culture irradiations in T-75 tissue culture flasks.

3.1.5 Flow cytometry analysis of cell cycle and viability

Prior to RS analysis, the cell cycle distribution and viability of each culture was measured using flow cytometry (recall description in section 1.2.1), following established protocols [154, 155]. For cell cycle analysis, over 200,000 cells were extracted from a culture and fixed in 70% ethanol to permeabilize the cell membrane. The enzyme RNase A (Qiagen Inc., Mississauga, ON, Canada) was added to a concentration of 1 mg/mL in order to degrade cellular RNA. Propidium iodide (PI) (Sigma-Aldrich, Oakville, ON, Canada) was subsequently added to a final concentration of 50 μg/mL. After 30 minutes the suspension was centrifuged, re-suspended in a buffer (PBS + 1% FBS), and kept on ice until analysis. For flow cytometry collection, the relative DNA content of 100,000 cells was measured using a BD FACSCalibur Flow Cytometer (BD Biosciences, Mississauga, ON, Canada). Cell counts were recorded by the flow cytometer during sample acquisition. Relative fractions of cells in each phase of the cell cycle (G1, S, or G2) were determined using MatLab (Matlab, The Mathworks,

Natick, MA, USA) by performing a nonlinear least-squares fit to the measured data using functions representative of the expected distributions of DNA content for each cell cycle phase. G1 and G2 phase cells both produce Gaussian distributions of DNA content, and S phase cells produce a uniform distribution of DNA content between the central positions of the G1 and G2 distributions [19].

The cell membrane of a live cell does not allow PI to pass into the cell, meaning that flow cytometry of unfixed cells treated with PI can be used to test the viability of a cell culture by measuring the fraction of dead cells (positive PI signal) in a sample of cells. For culture viability assessment, over 100,000 cells were extracted from the culture and split into two equal parts. One part was stained with PI at a concentration of 5 μg/mL, and the other part was left untreated to serve as a control. 20,000 cells from each sample underwent flow cytometry analysis within 15 minutes of staining, and the fractions of live (no PI signal) and dead (positive PI signal) cells in the PI stained sample were determined with Matlab.

3.2 Raman spectroscopy

3.2.1 InVia Raman microscope

Raman analysis was performed on an inVia Raman microscope (Renishaw Inc., Illinois, IL USA) with a 100X dry objective (N.A. = 0.9) (Leica Microsystems, Wetzlar, Germany) and a 1200 lines/mm diffraction grating (figure 3.4). A 785 nm continuous wave diode laser (Renishaw) was used for sample excitation, providing a laser power density at the sample of $\sim$0.5 mW/μm^3. The laser power was measured using a Laser Power/Energy Meter (Coherent Inc., Santa Clara, CA, USA). The size of the sampling volume was measured to be $\sim$2 x 5 x 10 μm, using the 10-90% rise distance of the system edge response function for each axis [28]; these dimensions allow a single RS acquisition to represent the spectrum of a single cell ($\sim$10 μm in diameter), as further discussed in chapter 5. The spectral resolution of this configuration was measured to be 3 cm^{-1} at 520 cm^{-1}, using the full-width at half-maximum of a silicon line spectrum deconvolved with a Lorentzian function of the natural silicon linewidth [28, 156].

The laser and Raman light paths through the inVia Raman system are shown in figure 3.4. The excitation laser is first passed through a set of selectable neutral density filters, allowing for variable attenuation of the incoming beam. The shape

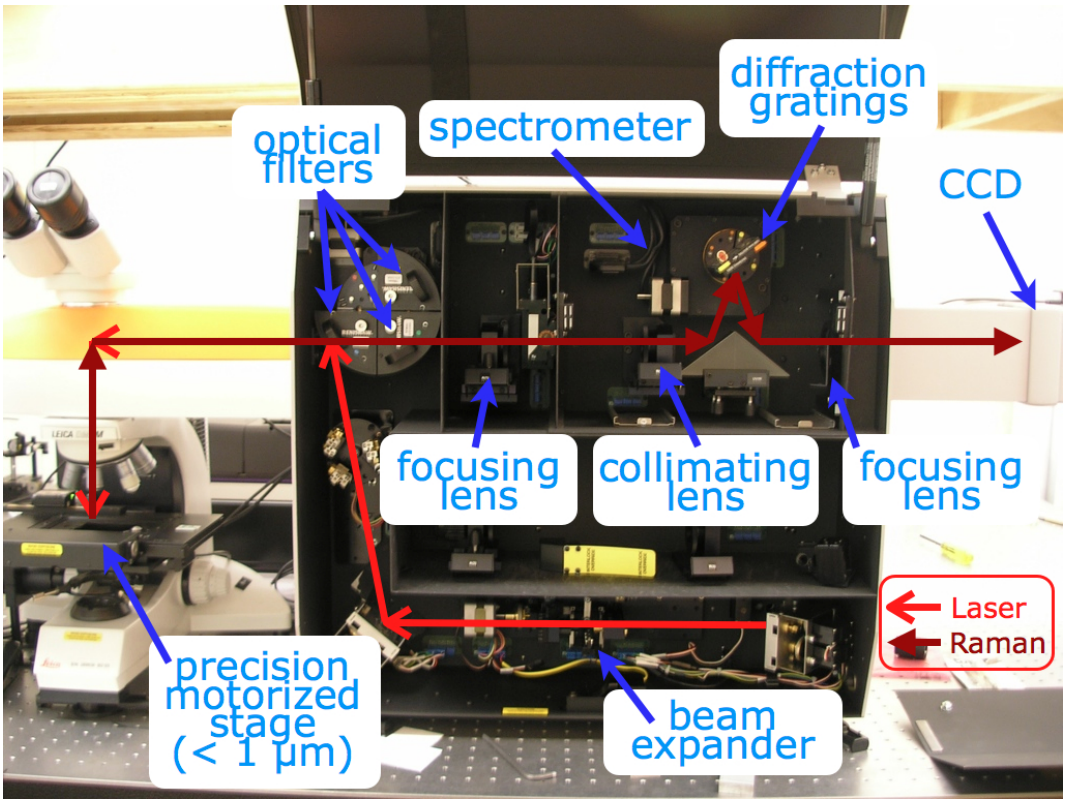


Figure 3.4: Light path through the inVia Raman microscopy system used in this work.

of the beam out of the laser is a $\sim 5 \times 2$ mm rectangle, which should be expanded to a circular beam to fill the back aperture of the 100x microscope objective (5 mm diameter). This is performed using a beam expander, consisting of a focusing objective followed by a collimating objective of lower magnification, each with a 5 mm diameter circular aperture. The resulting collimated beam exiting the beam expander is therefore circular with a 5 mm diameter. The laser beam is then reflected upwards by a mirror such that it is reflected off a 785 nm low-pass (LP) optical filter (Renishaw). The filter consists of hundreds of thin films deposited in such a way to reflect the 785 nm laser light, but pass light of any longer wavelength with a high transmission efficiency ($> 90\%$). The laser light is reflected off the LP filter into an upright microscope (Leica), and reflected downwards onto the back aperture of the microscope objective. The laser light is focused by the objective onto the sample, which is mounted on a linearly encoded motorized stepper stage with a minimum step size of 0.1 μm and repeatability of ± 0.2 μm (Prior Scientific Inc., Rockland, MA, USA). The isotropically scattered light, consisting of both Rayleigh and Raman

scattering, is collected and collimated by the microscope objective. The collected light reflects back along the incident beam path to the LP filter, where the Rayleigh light is reflected and the Raman light is transmitted through the filter. The transmitted Raman light is focused by a lens into the spectrometer chamber, where it follows the collimation, reflection and dispersion pattern shown in figure 2.4, before being re-focused along the dispersion axis (recall figure 2.5) of a 393 x 578 pixel RenCam CCD detector (Renishaw), thermoelectrically cooled to -70 °C to minimize thermal noise. All components of the inVia Raman system are controlled using the WiRE software package, version 3.0 (Renishaw).

For the inVia Raman system configuration used in this work, the dispersion axis of the CCD detector is 393 pixels wide, which corresponds to a spectral window 376 cm^{-1} wide when centred at 1000 cm^{-1}. To cover the two spectral windows used in this work (600-1800 cm^{-1} and 2700-3100 cm^{-1}), Renishaw's continuous scanning technique "SynchroScan" was used. This technique works by continually stepping the diffraction grating through the range of the desired extended spectral window while the accumulated charge at each pixel is shifted along the dispersion axis of the CCD in time with the rotation of the diffraction grating, such that the accumulated charge corresponding to a specific wavenumber is preserved as it is shifted along the CCD. The timing of the diffraction grating steps and the CCD pixel shifting is determined by the total acquisition time desired per pixel, such that each accumulated charge value (corresponding to one wavenumber) recorded at the far end of the dispersion axis will have accumulated for exactly the desired acquisition time [114]. The use of such continuous scanning techniques sets a lower limit on the desired acquisition time (10 seconds for the system used in this work) due to the physical and electronic constraints imposed by rotating the grating and shifting charges across the width of the CCD. It also requires additional time (equal to the acquisition time per wavenumber) to acquire the data, compared to the total time required using multiple acquisitions, since the first and last wavenumber regions in the spectra must both traverse the width of the CCD dispersion axis before being recorded. However, such continuous scanning techniques have two significant advantages over multiple acquisition techniques. Firstly, they eliminate the need for stitching separate spectra together, which can result in artifacts at the joining positions arising from variations in baseline intensity between acquisitions. Secondly, the shifting of charges across the CCD serves to average out any pixel-to-pixel variations across the detector, as every charge recorded for a given wavenumber has been accumulated from all of the pixel positions along

the CCD dispersion axis [114].

3.2.2 Single-cell spectral acquisition

Raman spectra were acquired from 20 individual cells from each sample, with cells chosen at random from the top layer of the cell pellet (figure 3.5). Spectra were collected at 30-second acquisitions per cell, for two different spectral windows: a low-wavenumber (LWN) window (600-1800 cm^{-1}) and a high-wavenumber (HWN) window (2700-3100 cm^{-1}). Spectra from each window were acquired in succession for each cell, using Renishaw's SynchroScan operation mode to cover each spectral window in a single acquisition.

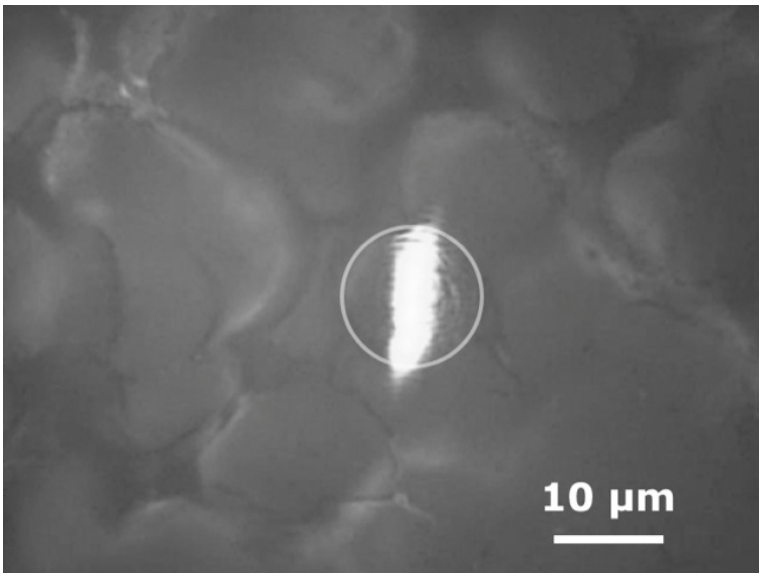


Figure 3.5: Optical image of a portion of a cell pellet, acquired with a 100X objective, with the 785 nm focused laser spot shown relative to the selected cell highlighted by the white circle.

All Raman spectra were acquired within 2 hours of transferring the cell pellet from the ice bath to the quartz disc. No spectral variations have been observed that correlate with the time of sample removal from the ice bath, suggesting that any effect of removing cells from the ice bath to an exposed environment has a negligible impact on the RS analysis within this 2 hour time interval.

Any spectrum obtained with excessive interference from cosmic rays or fluorescence was rejected ($\sim$1 in 20), and a replacement spectrum was obtained immediately from the same cell. The spectrum obtained from the same cell under repeated ac-

quisitions was highly reproducible, with only slight reductions in baseline intensity arising from laser-induced quenching of biological fluorescence. Any LWN window cell spectrum acquired resembling that of a dead cell [59] was rejected, and a new cell was selected. The acquisition of dead cell spectra was very infrequent, being dependent on the viability of the sample (measured in advance with flow cytometry). Most samples used in this work contained >95% live cells, and the threshold for proceeding to RS analysis was set at 75%. No correlations were found between the measured viability of harvested cultures and the results obtained from applying PCA to the RS data collected.

3.3 Spectral processing

The large number of spectra acquired in this work (80 to 240 per data set) necessitated the development and implementation of automated signal processing methods. Each spectrum from a given experimental data set was processed to (1) remove signals from cosmic rays, (2) increase the signal-to-noise ratio (SNR) via spectral smoothing, (3) estimate and subtract a baseline arising from the quartz substrate and biological fluorescence, and (4) normalize to the amount of biological material within the sampling volume. A description of the methods applied for each processing step is provided below (sections 3.3.1 - 3.3.4), ending with a description of principal component analysis (PCA) (section 3.3.5). All data processing and PCA was performed using Matlab.

3.3.1 Cosmic ray removal

Signals from cosmic rays (CRs) show up in Raman spectra as sharp spikes in the measured intensity, and typically span no more than a few pixels on the CCD. An automated CR removal program was developed, which was successful in removing CRs with peak intensities greater that approximately twice the average value of the spectrum in the local neighbourhood of the CR. This was done by iterating over each point in the spectrum, and if the value at a tested point was greater than twice the mean value of a 5-pt window adjacent to the tested value (sampled from the data already tested), then the value was replaced by a new value taken from its local neighbourhood. To ensure that the whole span of the CR was removed, the values of the points within ± 2 indices from the tested value were also replaced with new

values taken from their local neighbourhood. This algorithm processed ∼100-200 spectra per second, and was successful in removing ∼75% of the CRs detected during this work. Small CRs with peak values less than twice the average value of the local neighbourhood, and CRs with widths greater than a few pixels, were identified by visual inspection and removed manually via linear interpolation between points adjacent to the CR signal.

3.3.2 Two-point maximum entropy method smoothing

The two-point maximum entropy method (TPMEM) is a regularization method that has found utility for SNR enhancement of a variety of data types [157, 158], but has been particularly successful for use with Raman spectra [159, 160]. The basic theory of TPMEM is as follows. Two adjacent points in an array of data are considered to be a 2-point probability distribution. If the values of two adjacent points at indices i and $i+1$ are denoted x_i and x_{i+1}, then the fractional probability for each point can be written as

$$p = \frac{x_i}{x_i + x_{i+1}}, \qquad q = \frac{x_{i+1}}{x_i + x_{i+1}} \tag{3.1}$$

where $p + q = 1$. The two-point entropy function is defined as $-(p \ln p + q \ln q)$ [161]. TPMEM first generates the sum of the 2-point entropies, S, over all x_i. TPMEM then attempts to minimize the function $T = -S + \lambda\chi^2$, where λ is a Lagrange multiplier and χ^2 is the chi-square goodness-of-fit function between the recovered data and the original data. By minimizing T, TPMEM achieves a balance between smoothness of the recovered data (through S) and fidelity of the recovered data to the original data (through χ^2). Weighting between smoothness and fidelity is determined by choice of λ. A root-bisection algorithm is used to optimize λ subject to the expectation that $\chi^2 = X \cdot \mathrm{NP}$, where NP is the number of data points and X is a user-defined parameter that biases the recovered data to either increased smoothness ($X > 1$) or increased fidelity ($X < 1$). An example of applying TPMEM to a sample LWN window Raman spectra acquired from a cell, with X varying from 0.25 to 1.25, is shown in figure 3.6. For the cell spectra acquired in this work X was set to 0.25 for the LWN window and 0.75 for the HWN window; these values were found to provide a modest level of smoothing, without sacrificing fidelity to the original intensity of the Raman signals. The TPMEM algorithm used in this work processed ∼50-100

spectra per second.

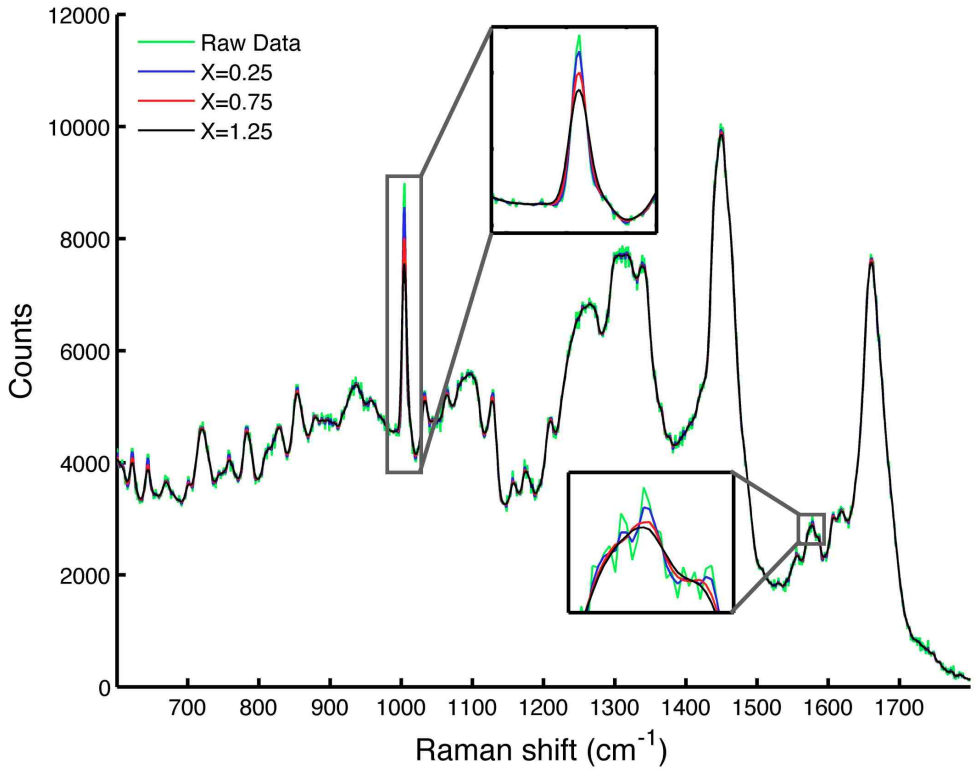


Figure 3.6: Example of applying the two-point maximum entropy method (TPMEM) for smoothing of a raw Raman spectrum of a cell, for the LWN spectral window, with the user-defined parameter X varying from 0.25 to 1.25.

3.3.3 Baseline estimation

Baseline (BL) estimation and removal is necessary for comparing Raman signals between different spectra, as it is the Raman peak height, or area, above the BL that provides useful information about Raman scattering within the sample. An effective and robust BL estimation method is critical for this work in order to remove sources of variability arising from varying levels of fluorescence or quartz substrate contamination in a multi-spectra data set (further discussed in chapters 4 and 5). In this work, different BL estimation methods were applied for the LWN and HWN spectral windows to address the different BL structure in each window. The BL in the LWN spectral window arises from the combined contributions from biological fluorescence, the quartz substrate, water / culture media and optical components in the Raman

system, whereas the BL in the HWN spectral window arises only from very weak contributions from biological fluorescence and water / culture media.

Low-wavenumber spectral window

For the LWN spectral window, the algorithm used in this work is a modified version of the signal removal method (SRM) as described by Schulze *et. al.* [122]. An initial crude estimate of the BL is generated by applying a first-order Savitsky-Golay (SG) filter [162] to the original data, using a SG window size (W) of 3-7% of the total range of the data (1304 points). This initial estimate provides a threshold that separates "signal" (data above the threshold) from "baseline" (data below the threshold). The original data is then modified by replacing any data above the threshold with the value of the SG filter BL estimate at that point (*i.e.*, the signal is removed). The SG filter is then applied to the modified data set, providing a second estimate of the BL. The signal removal step is repeated using the SG filter output as the new threshold, and the SG filter is applied again to the modified data set. The process is repeated until a stopping criterion is reached, or a fixed number of iterations is completed. In this work, fixing the number of iterations to 20 proved to be sufficient, as further iterations did not change the BL. An example of applying the SRM used in this work to a LWN window TMPEM smoothed cell spectrum is shown in figure 3.7, where the SG filter BL estimate and the SRM modified data are shown for various iterations of the algorithm. The completion time for this algorithm was ∼1 second per spectra.

The SRM method was chosen for this work due to the mixture of sharp and broad features throughout the LWN spectral window. As such, it is desirable to use a highly conformal BL estimate around regions of possible quartz contamination (∼800 cm^{-1} and ∼1050 cm^{-1}) that underly sharp Raman features, and a broader BL estimate from 1200 - 1800 cm^{-1}, where many overlapping peaks give rise to broad Raman features. The most important parameter in the SRM algorithm was found to be the window size of the SG filter. Smaller window sizes (*e.g.*, 3% of the data range) allow more variation in the baseline estimate, whereas larger window sizes (*e.g.*, 7%) yield less variation. This effect on the final SRM BL estimate, for different single window sizes (W), is shown in figure 3.8.

As shown in chapter 4, a single window size is not optimal for BL estimation around all the features of the LWN spectral window (*e.g.*, figure 3.8). In this work a SG window size of 3% was used for the first half of the LWN spectral window

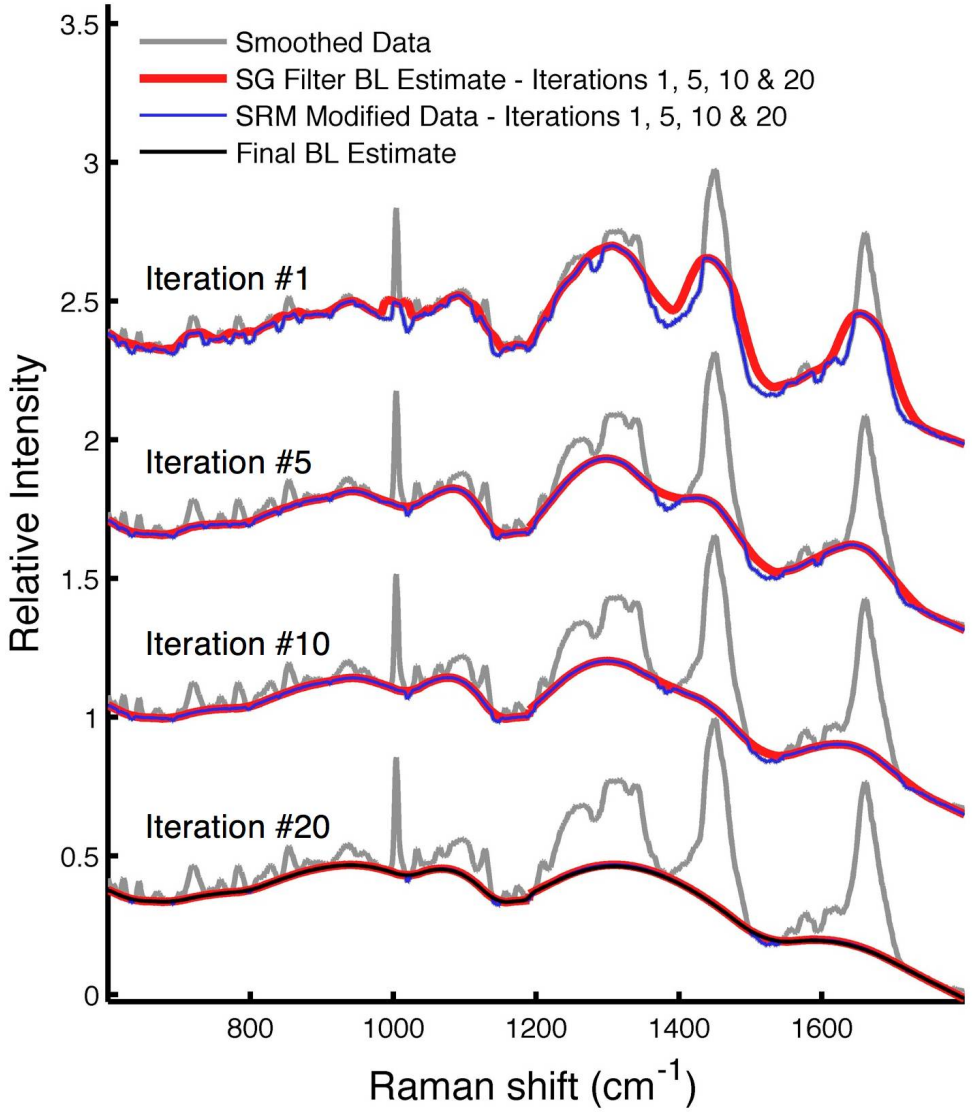


Figure 3.7: Example of applying the signal removal method (SRM) used in this work for baseline (BL) estimation of a smoothed cell spectrum, showing the result of the Savitsky-Golay (SG) filter BL estimate, and the signal removal method (SRM) modified data, for various iterations of the algorithm.

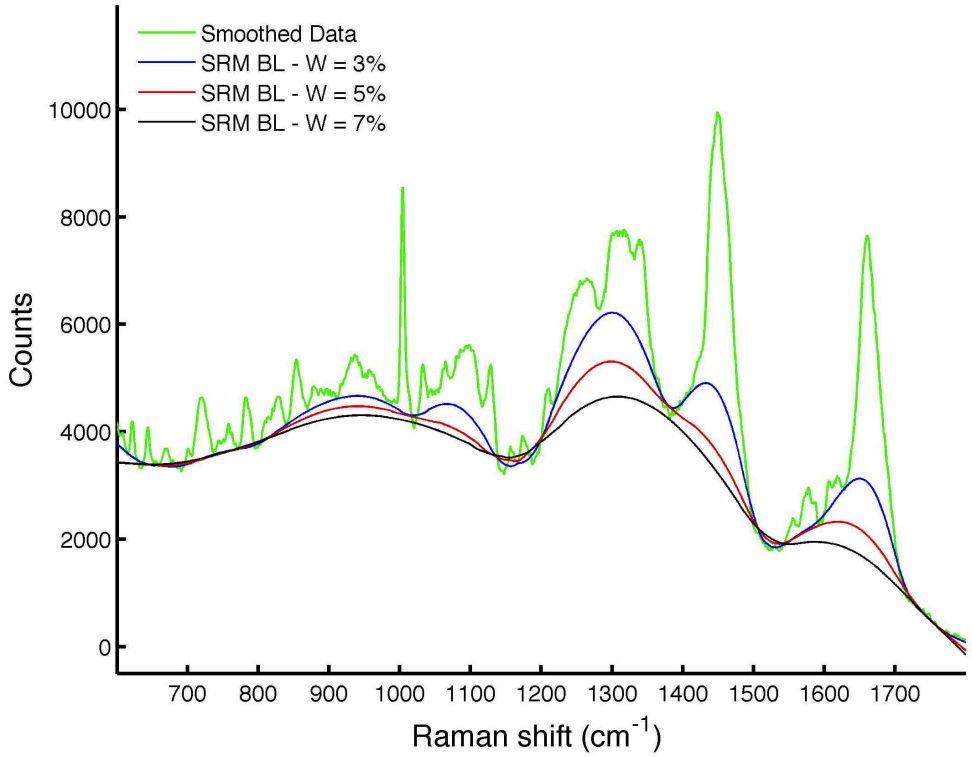


Figure 3.8: Example of applying the signal removal method (SRM) for baseline (BL) estimation of a smoothed cell spectrum, showing the effect of varying the SG filter window size (W) from 3% to 7% of the total range of the data

(600 - 1190 cm^{-1}) and a SG window size of 7% was used for the second half of the LWN spectral window (1191 - 1800 cm^{-1}). The two BLs were concatenated, and the discontinuity at 1190 cm^{-1} was removed by adding a small amount of constant amplitude oscillatory noise to the concatenated BL (noise amplitude was set at 5% of the mean value of the BL), then applying a high level of TPMEM smoothing (X=1.5) to the noisy BL. The output of this method is shown in figure 3.9, along with the spectrum obtained after subtraction of the final BL estimate.

High-wavenumber spectral window

For the HWN spectral window, the algorithm used in this work employs a three-point linearly interpolated (TPLI) method for baseline (BL) estimation. This method was chosen since there is very little variability in baseline structure in this spectral window, and the Raman signals are very broad, spanning almost the whole width of the spectral window. Three small regions (each 5 cm^{-1} wide) of the HWN window

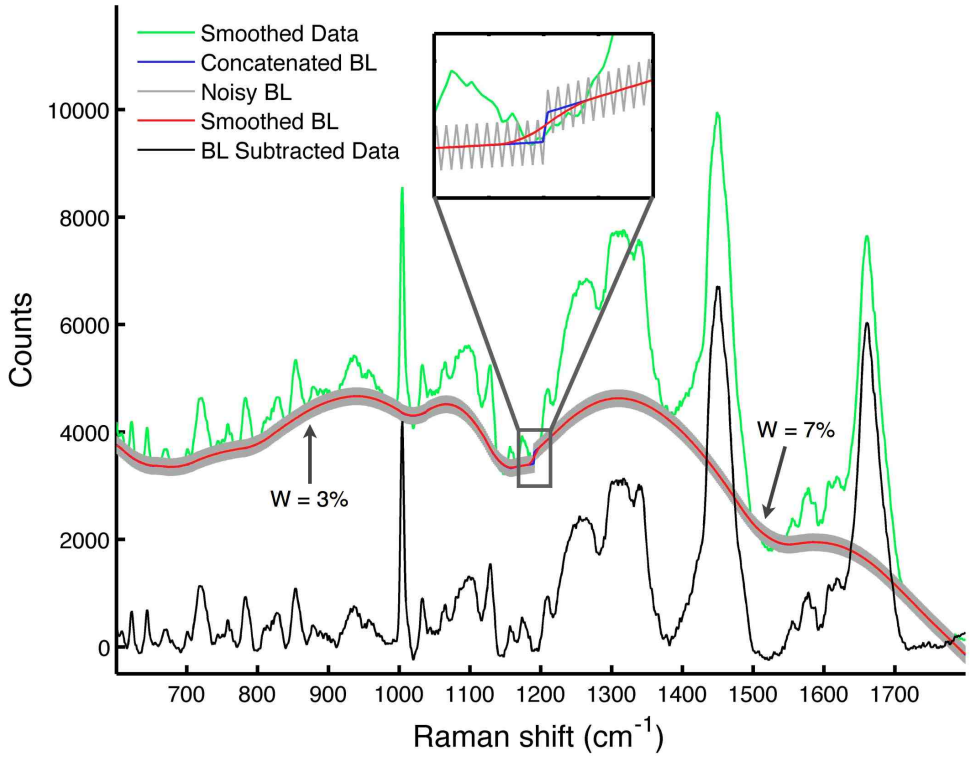


Figure 3.9: Example of applying the signal removal method (SRM) for baseline (BL) estimation for a low-wavenumber (LWN) spectral window cell spectrum, showing the method of using different window sizes (3% and 7%) for each half of the LWN window then concatenating the two BLs and removing the discontinuity at 1190 cm^{-1} by adding noise to the concatenated BL and smoothing with TPMEM. The resulting BL subtracted spectrum is also shown.

spectrum were selected, which were found to consistently represent areas without Raman signals (*i.e.*, only contributions from background signals). The minimum value of each region was chosen as a BL point, and linear interpolation was then performed between the three selected points and the first and last points in the spectrum. An example of this method applied to a HWN cell spectrum is shown in figure 3.10, showing the three regions designated as background and the spectrum obtained after subtraction of the BL estimate.

3.3.4 Spectral normalization

After BL estimation and subtraction, the principal remaining source of variability between spectra is the overall intensity of the Raman features, which arises from the variable amount of biological material within the sampling volume and from

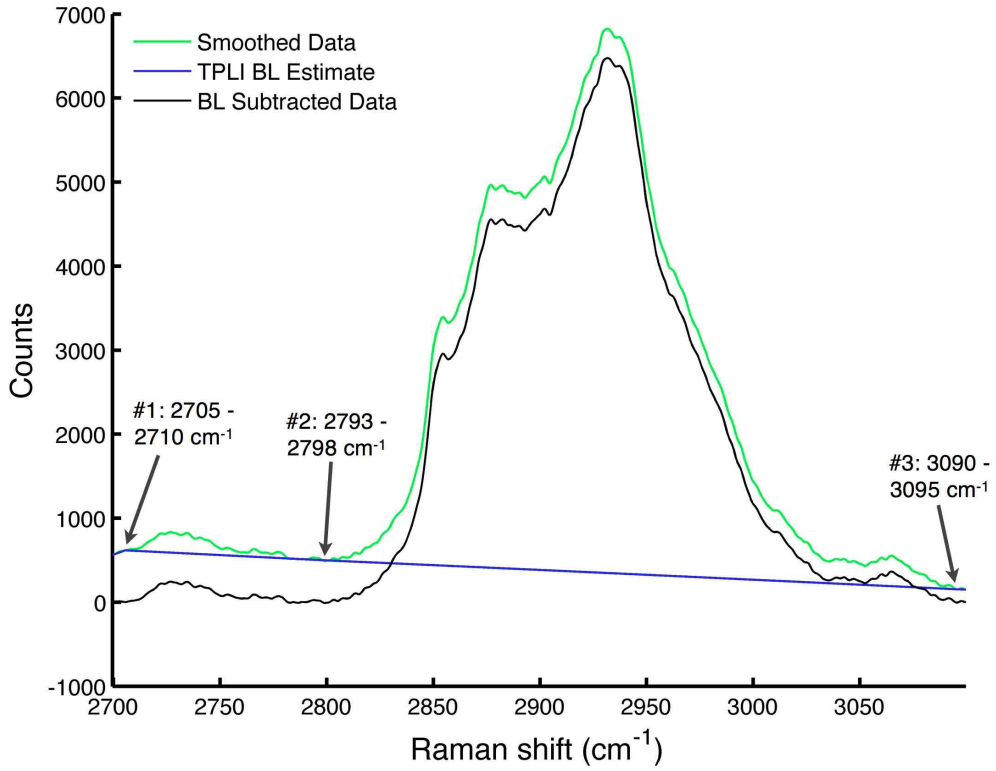


Figure 3.10: Example of applying the three-point linearly interpolated (TPLI) method for baseline (BL) estimation for a high-wavenumber (HWN) spectral window cell spectrum, showing the three small regions designated as background from which the BL points were chosen. The resulting BL subtracted spectrum is also shown.

fluctuations in laser power. In this work, this variability predominantly arises from slightly different physical shapes and orientations of each cell in the cell pellet, relative to the Raman sampling volume (see figure 3.5). To remove this source of variability, each spectrum was normalized to the total area under the BL subtracted spectrum. Other authors have addressed this issue of intensity variability by normalizing to the area under the CH deformation peak at 1450 cm^{-1}, thought to be proportional to the total amount of biological material within the sampling volume [59, 81]. In this work, it is found that the area of the CH deformation peak can vary independently of other Raman peaks, and therefore may not be suitable as a normalization peak in all cases. For example, in chapter 5 it is shown that one of the most significantly varying Raman signals between cells arises from CH_2 deformation in lipids at 1438 cm^{-1}, which affects the area of the CH deformation peak as well due to its close spectral proximity. These considerations are further addressed in chapter 4. In this work it was found that normalizing to the total area under the BL subtracted spectrum is

suitable for both the LWN and HWN spectral windows, and is demonstrated in figure 3.11.

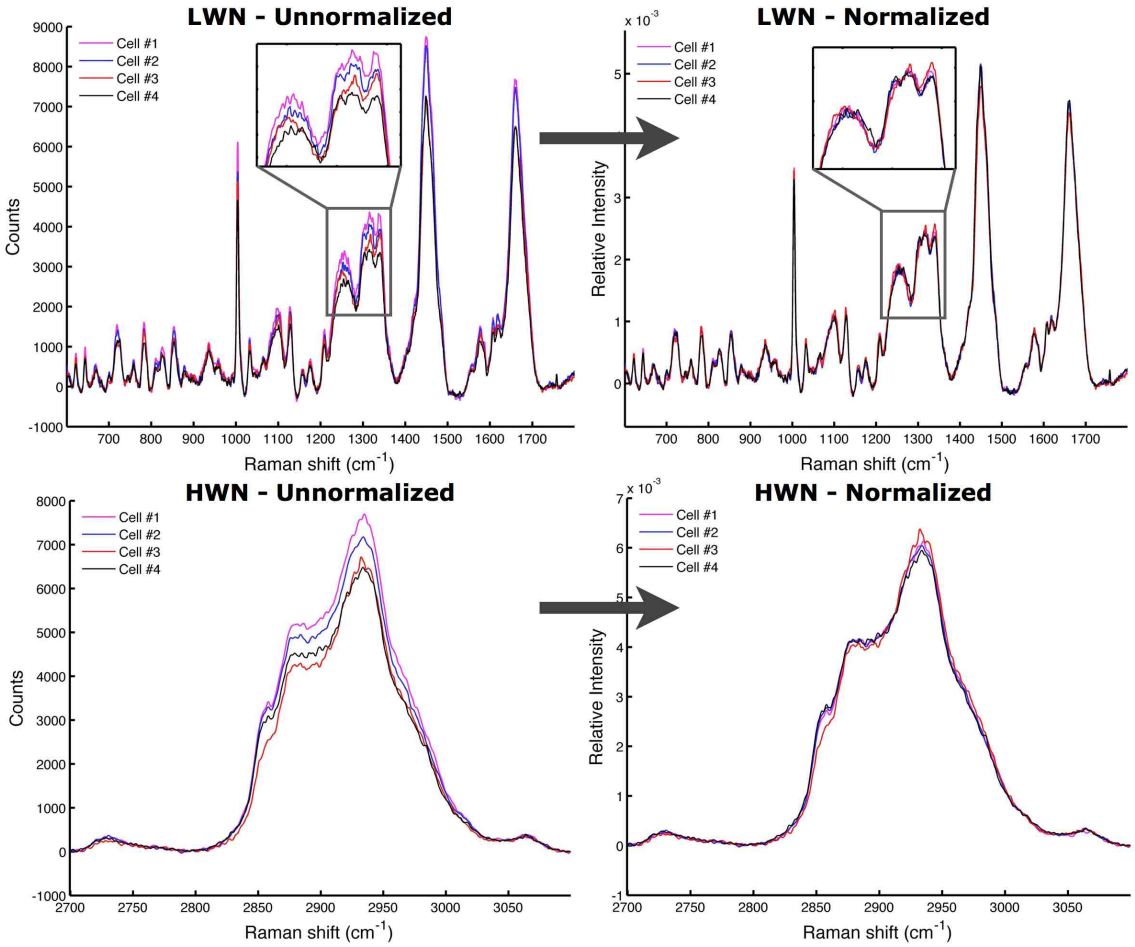


Figure 3.11: Example of applying normalization to four baseline (BL) subtracted cell spectra, using the total area under the BL subtracted spectrum, for both the low-wavenumber (LWN) and high-wavenumber (HWN) spectral windows.

3.3.5 Principal component analysis

PCA utilizes a mathematical transform that reveals the internal structure of a data set in a way which best isolates the independent sources of variance in the data. A brief overview of the theoretical aspects of PCA that are relevant to this work is provided below. PCA was performed in this work using standard algorithms (Matlab).

Given an input data set of observations $\mathbf{O}$ of size $m \times p$, where m is the number of observations (*i.e.*, number of spectra) and p is the number of variables (*i.e.*, points

per spectra), PCA first subtracts off the mean of all observations (O_{avg}) from each individual observation, then calculates the sample covariance matrix $\mathbf{S}$ of the mean-subtracted data set $\mathbf{M} = \mathbf{O} - O_{avg}$ [163]. PCA then calculates the first $m-1$ eigenvectors of $\mathbf{S}$, which are here called the "components" and denoted $C_1 - C_{m-1}$, each vectors of length p. These eigenvectors of $\mathbf{S}$ are also sometimes referred to as "coefficient" or "loading" vectors [163]. The components form a complete orthonormal basis set of the mean-subtracted data, such that ${C_n}' \cdot C_m = 0 \; \forall \; n \neq m$, and ${C_n}' \cdot C_m = 1 \; \forall \; n = m$ (here $'$ denotes the vector transpose operation). The eigenvalues of $\mathbf{S}$ corresponding to each component are also calculated, and each eigenvalue is equal to the total variance explained by the corresponding component [163]. As such, PCA orders the components, $C_1 - C_{m-1}$, by the amount of variance each component explains in the mean-subtracted data set $\mathbf{M}$. As an example, figure 3.12 shows the 1st, 2nd, 20th, 50th and 239th (last) PCA components from a 240 spectra LWN window data set used in this work. As shown, the percent variance explained decreases with each component.

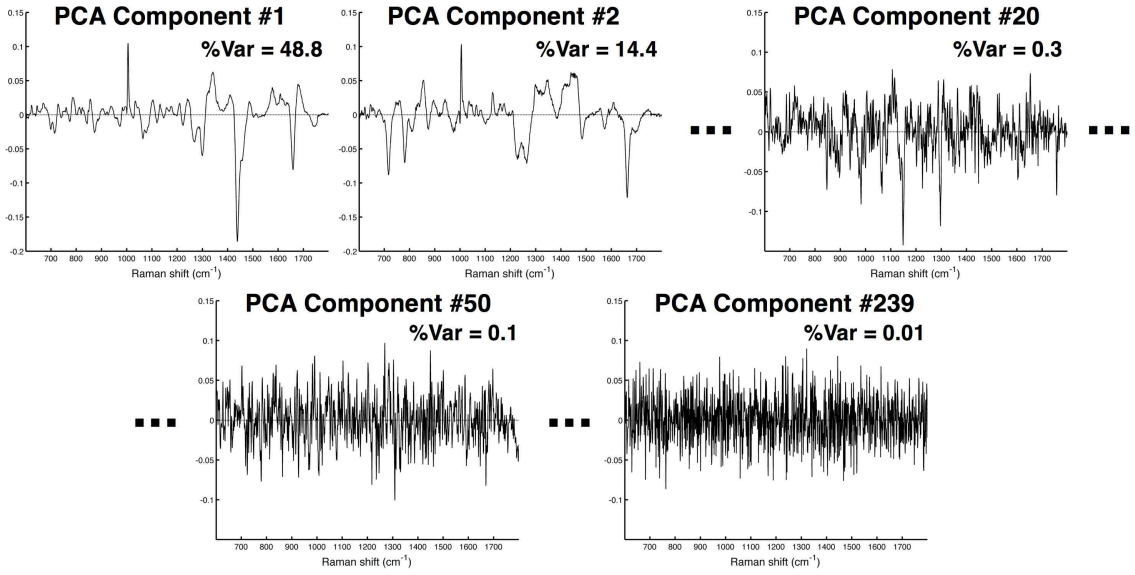


Figure 3.12: Examples of PCA components from a 240 spectra LWN window data set used in this work. The PCA components are ordered by decreasing percent variance explained in the original mean-subtracted data set.

The first two components in figure 3.12 explain the majority (63.2%) of the variance in the data set and contain clearly defined positive and negative features, indicating correlated sources of variability at the corresponding wavenumbers of the mean-subtracted data set. The 20th component explains 0.3% of the total variance

and contains few defined features, being dominated by random noise. The 50^{th} and 239^{th} components have no visible features, and the percent variances are reduced to 0.1% and 0.01% respectively. Since most of the non-random sources of variability in the data set are accounted for by the first $n << m-1$ components (the number of meaningful components n depends on the data set), PCA effectively reduces the dimensionality of the data set to a more manageable size (*e.g.*, the set of components each explaining $> 2\%$ of the total variance) for analysis and interpretation of the primary sources of variability in the data set.

To connect the correlated sources of variability (*i.e.*, the components) to the spectra in the original data set, each original observation (*i.e.*, each spectrum in the original data set) is assigned a score ($s_{1,i} - s_{m-1,i}$) for each component, calculated by projecting each component C_j onto $\mathbf{M}$ such that

$$s_{j,i} = C_j{}' \cdot M_i \tag{3.2}$$

where M_i denotes the i^{th} mean-subtracted original observation, and $j = 1 \ldots m-1$. Each score determines how much each component contributes to the variance of the original observation (O_i) from the average observation (O_{avg}). Each original observation O_i can therefore be recovered by the following operation:

$$O_i = O_{avg} + s_{1,i} \cdot C_1 + s_{2,i} \cdot C_2 + ... + s_{m-1,i} \cdot C_{m-1} \tag{3.3}$$

The PCA components (C_j) provide direct observation of the dominant sources of variability within a data set, whereas the PCA scores ($s_{j,i}$) show how much each observation in the data set is responsible for a given source of variability. Analysis of the PCA scores for a large spectral data set provides insight into which spectra in the data set are causing the variability, and allows for the detection of trends between samples (or spectra) in the experiment. A schematic of the PCA score plots for the first and second PCA components, from a fictitious experiment with four samples of 20 spectra each, is shown in figure 3.13. In this work, the spectrum number (i) is plotted against the PCA scores for a certain component. By equation 3.3, any spectrum assigned a higher (*i.e.*, more positive) PCA score for a given component contains a proportionately higher amount of the positive features and a lower amount of the negative features from that component, in the original spectrum. It should be noted that the positive or negative nature of the features is purely arbitrary and

only holds meaning with respect to the sign of the corresponding PCA scores. Any component can be reflected about zero, with a corresponding change of sign for all scores for that component, without altering the results of the PCA transformation.

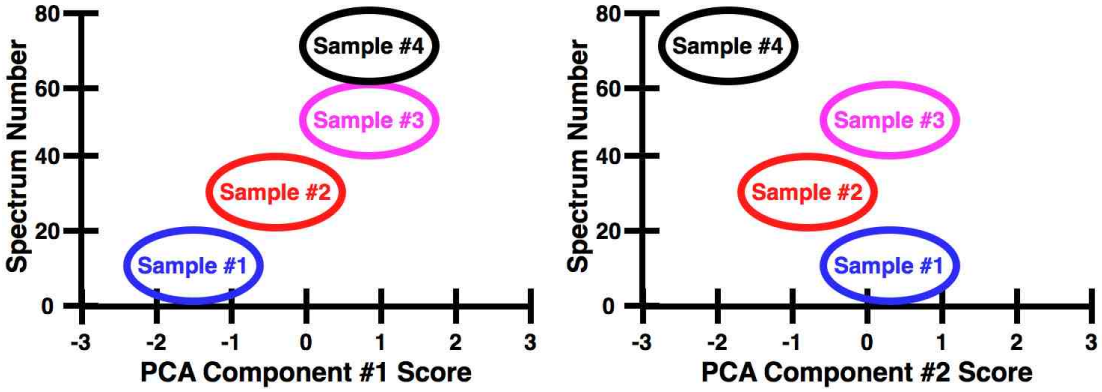


Figure 3.13: Schematic of the PCA score plots for the first and second PCA components, from a fictitious experiment with four samples of 20 spectra each.

Since multiple spectra are collected per sample, the PCA scores are grouped according to their sample of origin (figure 3.13), often indicating differences in some controlled variable in the experiment (*e.g.*, time, dose, treated vs. control). Shifts in the PCA score distributions along the X-axis, for one or more samples relative to the other samples, indicate that the corresponding PCA component describes spectral variability arising predominantly from spectral differences between the samples. For the results of the fictitious experiment depicted in figure 3.13, the PCA scores for the first PCA component indicate spectral variability arising predominantly from spectral differences in the first two samples relative to samples 3 and 4. The increase in the PCA scores for sample 2 (and 3 and 4) relative to sample 1 corresponds to increased amounts of the positive features relative to the negative features in the first PCA component, in the original spectra of sample 2 (and 3 and 4) as compared to sample 1. The PCA scores for the second PCA component indicate an independent source of spectral variability that arises predominantly from samples 2 and 4. If samples 2 and 4 are known to differ from samples 1 and 3 only according to some controlled variable in the experiment, the PCA scores serve to identify a source of spectral variability correlating with that experimental variable. The spectral features responsible for the variability in the original data set are then identified by the positive and negative features in the corresponding PCA component.

The statistical significance of an observed shift in the PCA score distribution

between two samples can be calculated using a two-sample t-test, which computes the probability that two samples are drawn from the same distribution [164]. This probability is called the "p-value". If the p-value is at or below a certain threshold, typically set at 0.05 (*i.e.*, 5%), then the two distributions are considered statistically different [164]. In this work, the difference between two PCA score distributions is evaluated using a two tailed t-test, without assuming equal variances between samples, computed using standard algorithms (Matlab).

A side application of PCA utilizes equation 3.3 to regenerate the original set of spectra, $\mathbf{O}$, without including one or more components. For example, regenerating the original spectra using only the first $n << m-1$ components effectively removes most of the random sources of variability (*e.g.*, spectral artifacts, noise) from the data set. In this way one can also filter out other undesirable components, such as a component arising from a sample substrate, as long as the unwanted contributions are isolated into a single component after the PCA calculation. For certain data sets in this work, included in chapter 5, excess spectral variability arising from the quartz substrate was easily identifiable in a single PCA component; the quartz component was therefore removed and the PCA calculation was repeated on the filtered set of spectra. This action does not affect the other components (or scores), only redistributes the variance explained by the excluded component among the remaining components.

Chapter 4

Results & Discussion I: Validation of Analysis Methods

4.1 Introduction

The purpose of this chapter is to validate the combination of data processing techniques used in this work (described in chapter 3, section 3.3) for the analysis of Raman spectra obtained from single human cells. A representative set of raw data is used to demonstrate the effect of varying certain processing parameters on the final outcomes of the data analysis. Comparisons of the outcomes of the sample data set are performed in two ways: (1) by inspection of two processed spectra representing typical levels of differences between spectra across the sample data set, and (2) by inspection of the PCA components and PCA scores obtained from the entire set of processed spectra.

The data set chosen for this analysis is from an irradiation experiment performed with the DU145 cell line, which is presented and discussed in full in chapter 6 and included in the results of chapter 7. There are 240 cell spectra in this data set, irradiated and collected for RS acquisition as described in section 3.1.4. This data set was chosen because of the clear trends of the PCA scores for both the first and second PCA components, providing an opportunity for easy inspection of the effects of varying the data processing parameters on the data outcomes. Only the LWN spectral window data is discussed, as the HWN window data is much less complicated and is therefore much less sensitive to variations in the data processing parameters.

The data processing steps that have the most potential for affecting the outcomes

of the data are: (1) spectral smoothing, (2) baseline estimation, and (3) spectral normalization. The effects of altering the parameters in these steps, relative to the parameters used in this work, are each presented in turn (sections 4.2 - 4.4), followed by a general discussion of the effect of varying the data processing parameters with respect to the aims of the RS studies in this work (section 4.5). Molecular and biological interpretations of the data outcomes are not discussed in this chapter, with the exception of section 4.4 where the choice of normalization method has an important effect on the molecular interpretations of the data outcomes.

4.2 Spectral smoothing

The only user-defined parameter in the TPMEM smoothing method is the X value, which (recall section 3.3.2) biases the recovered data to either increased smoothness ($X > 1$) or increased fidelity to the original data ($X < 1$). High levels of smoothing may degrade the accuracy of Raman signal intensities, especially near sharp peaks and "valleys" in the spectra, as shown in figure 3.6 for $X = 0.75$ and $X = 1.25$. Not applying any smoothing at all is the most conservative approach, but will inevitably leave more random noise in the data, possibly reducing the ability to detect subtle spectral features. The effects of no smoothing ($X = 0$) and smoothing using $X = 1.25$ are presented below, relative to smoothing with $X = 0.25$.

4.2.1 Effect on processed spectra

The effect of varying the TPMEM smoothing parameter X, from $X = 0$ (no smoothing) to $X = 0.25$ to $X = 1.25$, on two processed cell spectra is shown in figure 4.1, with a region of interest (ROI) highlighted to demonstrate the differences in random noise between the two pairs of spectra for each X parameter. From $X = 0$ to $X = 0.25$ (figure 4.1a and 4.1b), no loss of fidelity to the original data is apparent within the ROI, although there is a very slight decrease in peak area ($< 0.5\%$ for both spectra) from 1000-1009 cm^{-1}, the sharpest feature in the LWN spectral window. None of the other Raman signals are similarly affected. However, from $X = 0$ to $X = 1.25$ (figure 4.1a and 4.1c), there is a definite loss of fidelity to the sharp Raman features between 600 and 1200 cm^{-1}, most evident by the $\sim$6% decrease in peak area from 1000-1009 cm^{-1}, for both cell spectra. These losses of peak fidelity with increased smoothing are indicated by the arrows in figures 4.1a-c, which are fixed at the position of original

($X = 0$) intensity for the 1000-1009 cm^{-1} peak.

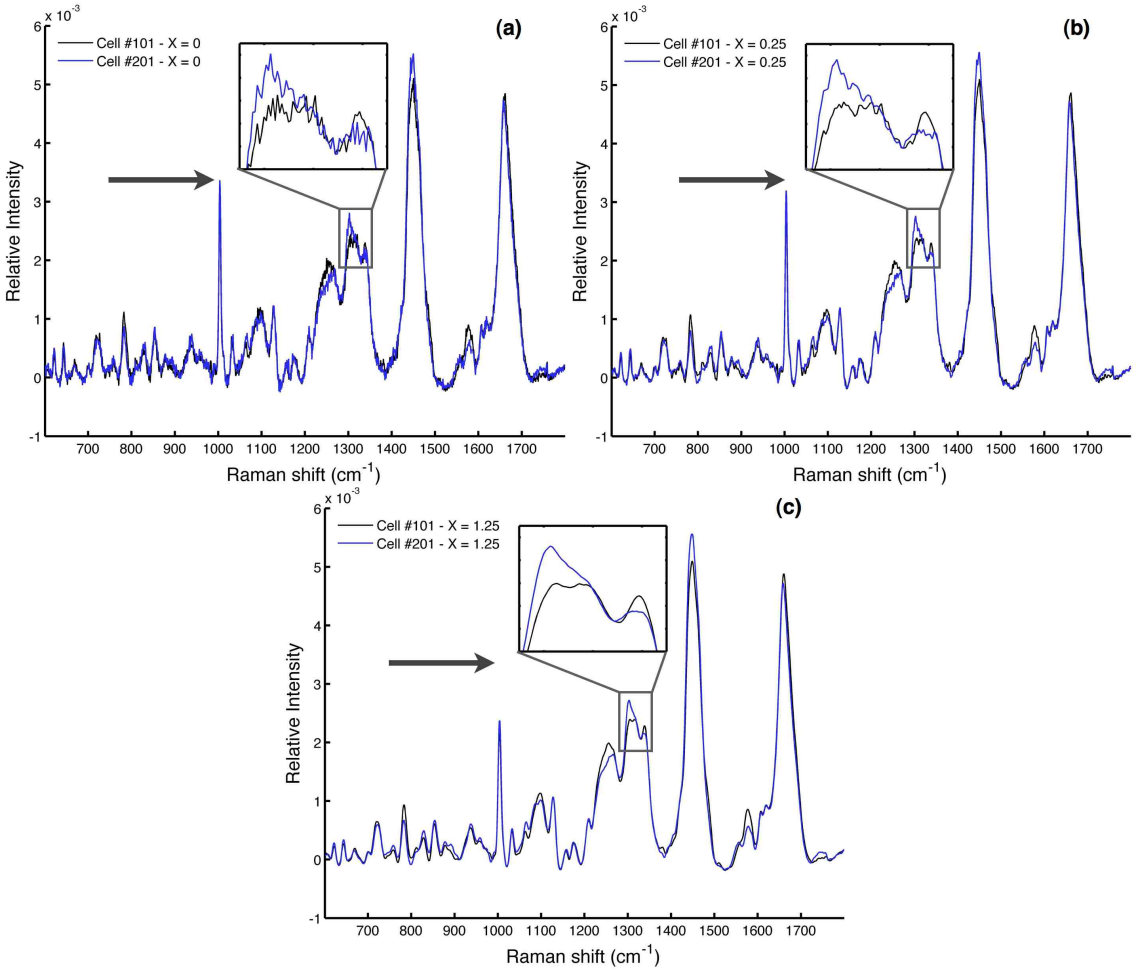


Figure 4.1: Effect of varying the TPMEM smoothing parameter X, from (a) $X = 0$ (no smoothing) to (b) $X = 0.25$ to (c) $X = 1.25$, on two processed DU145 cell spectra, with a region of interest highlighted to demonstrate the differences in random noise between the two pairs of spectra for each X parameter. The reduction in peak fidelity with increased smoothing is indicated by the arrows fixed at the position of original ($X = 0$) intensity for the 1000-1009 cm^{-1} peak

4.2.2 Effect on PCA components and PCA scores

For the 240 cell spectra in the sample DU145 data set, the first and second PCA components represent the most significant sources of independent spectral variability in the data set. The positive and negative features in the PCA components arise from variability in the intensity of the corresponding Raman peaks in the original data set

(in chapters 5-7, a molecular origin is assigned to each feature in the components). Recall, by equation 3.3, that any spectrum with a higher (*i.e.*, more positive) PCA score for a given component contains a proportionately higher amount of the positive features and a lower amount of the negative features from that component, in the original spectrum.

PCA components

The effect of varying the TPMEM smoothing parameter on the first two PCA components of the sample DU145 data set is shown in figure 4.2. The positive and negative features of both PCA components are accurately reproduced for all X values, with only a slight loss in the ability to resolve adjacent features for $X = 1.25$. There is also a slight reduction in the intensity of the sharp feature at 1004 cm^{-1} for $X = 1.25$, corresponding with the loss in fidelity to that peak observed in the processed spectra (figure 4.1c). The most significant effect of varying the TPMEM smoothing parameter on the first two PCA components is the increase in the percent variance explained by each component as the level of smoothing is increased. This increase is a direct result of the decrease in random noise in the data set with increased smoothing, as decreased random noise results in an decrease in the percent variance explained by the remainder of the PCA components, most of which account for random sources of variability in the data set, such as noise (recall figure 3.12).

PCA scores

The PCA scores corresponding to the first two PCA components determine how much of the variability explained by each component (figure 4.2) is expressed in each of the 240 cell spectra in the sample data set. The effect of varying the TPMEM smoothing parameter on the PCA scores for the first two PCA components of the sample DU145 data set is shown in figure 4.3. Only very small changes in the individual PCA score values are observed. The average absolute shift in the score values between $X = 0$ and $X = 1.25$ is $\sim 0.05 \times 10^{-3}$ for both components, only 4% and 9% of the average standard deviation of the sample PCA score distributions for the first and second components, respectively. The observed trends of the PCA scores, the average standard deviation of the sample PCA score distributions, and the levels of significance for the differences between PCA score distributions (*i.e.*, the p-values), are not affected by these variations in the level of TPMEM spectral smoothing.

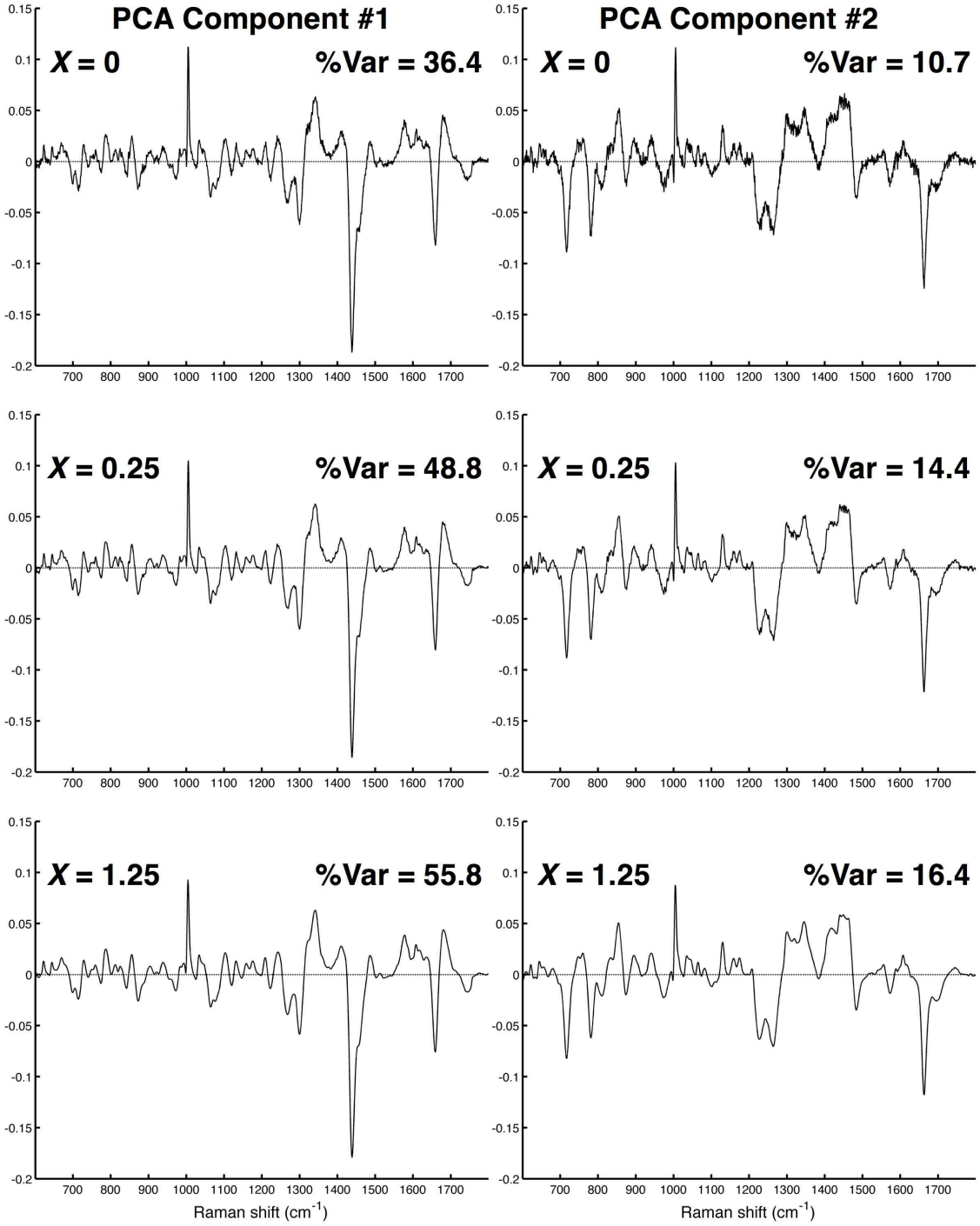


Figure 4.2: Effect of varying the TPMEM smoothing parameter X, from $X = 0$ (no smoothing) to $X = 0.25$ to $X = 1.25$, on the first two PCA components of the sample DU145 data set.

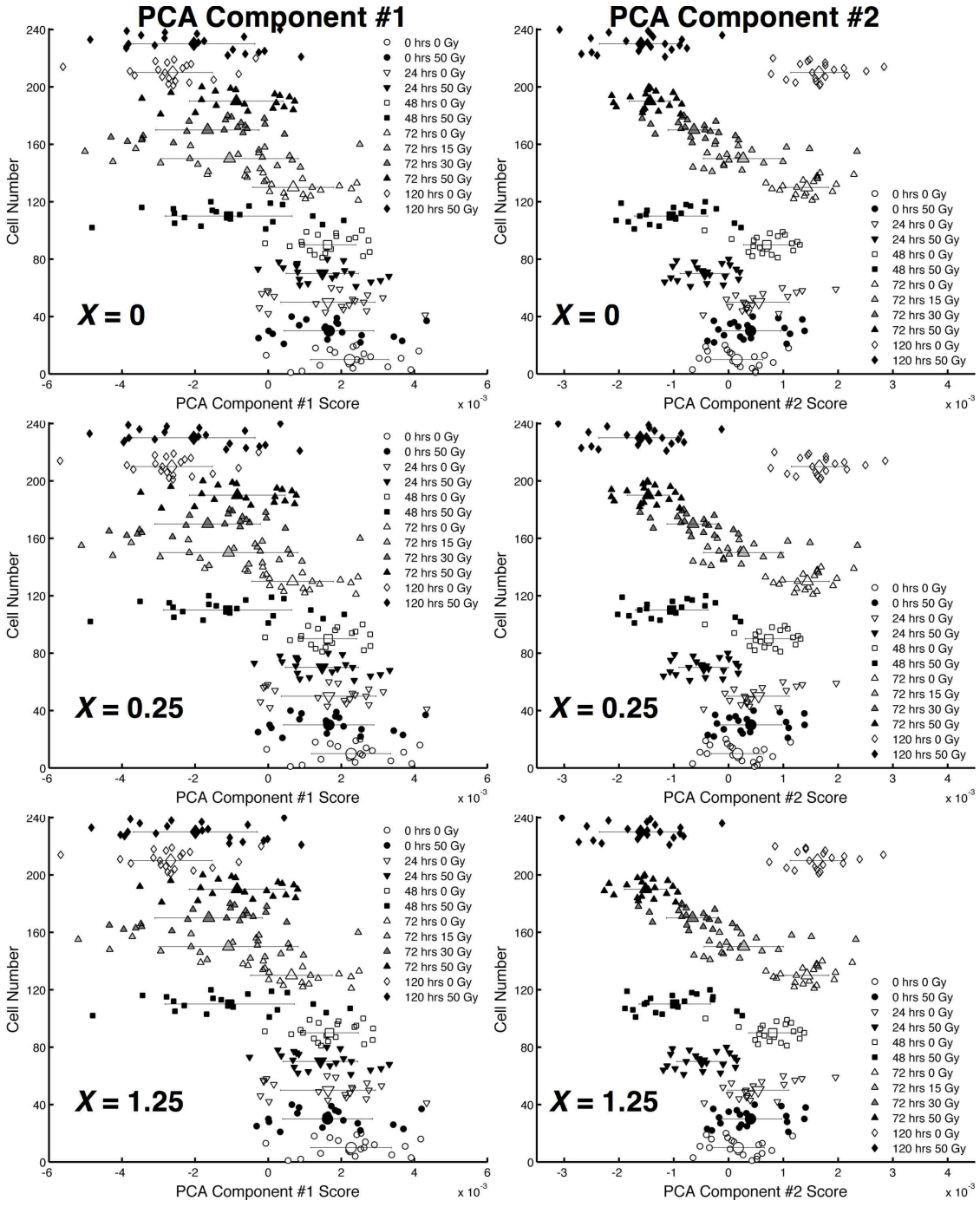


Figure 4.3: Effect of varying the TPMEM smoothing parameter X, from $X = 0$ (no smoothing) to $X = 0.25$ to $X = 1.25$, on the PCA scores for the first two PCA components of the sample DU145 data set. Different markers categorize all 240 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

4.3 Baseline estimation

For the LWN spectral window data in this work, the signal removal method (SRM) for baseline (BL) estimation is used with different Savitsky-Golay (SG) filter window sizes (W=3% and W=7%) for each half of the spectral window (described in section 3.3.3). This application of the SRM method, although similar to both the SRM method described by Schulze *et. al.* [122] and the "modified polyfit" method described by Lieber *et. al.* [165], is unique due to the use of different SG window sizes, and therefore different BL estimation parameters, for a single spectral window. The effects of, alternatively, using the SRM with single window sizes of 3%, 5% and 7%, for the entire LWN spectral window for the DU145 sample data set, are presented below.

4.3.1 Effect on processed spectra

Recall that the differences between the final BL estimates for the SRM with single window sizes of 3%, 5% and 7% are shown in figure 3.8 for a representative spectrum. As the window size is reduced the BL estimate becomes more conformal to the "valleys" of the data, which may or may not represent regions without contributions from Raman signals, depending on the amount of overlap of the Raman features in the vicinity. However, a less conformal BL estimate may not accurately correct for variable levels of fluorescence between spectra. The effect of using single window sizes of 3%, 5% and 7% on two processed cell spectra is shown in figures 4.4a to 4.4c, with a ROI (800-920 cm $^{-1}$) highlighted to show the variability in the subtracted BL level between the two spectra in a region of inherently variable fluorescence between spectra. The differences in the BL subtracted spectrum for a single cell using these window sizes is shown in figure 4.4d. The difference in overall relative intensity using different window sizes is the result of the difference in total area under the spectrum after BL subtraction, which is used for spectral normalization.

For the ROI highlighted in figures 4.4a to 4.4c, the separation between the subtracted BL levels between the two spectra is the smallest for W=3% and becomes larger for W=5% and W=7%. However, for wavenumbers greater than ∼1200 cm^{-1} there are no variations in the BL levels between the two spectra for each window size (although there are large differences in the BL levels for different window sizes, as seen in figure 4.4d). Since the Raman features in the 1200 - 1800 cm^{-1} region are predominantly broad features with overlapping contributions from multiple Raman signals, smaller window sizes inevitably remove more Raman signal intensity from

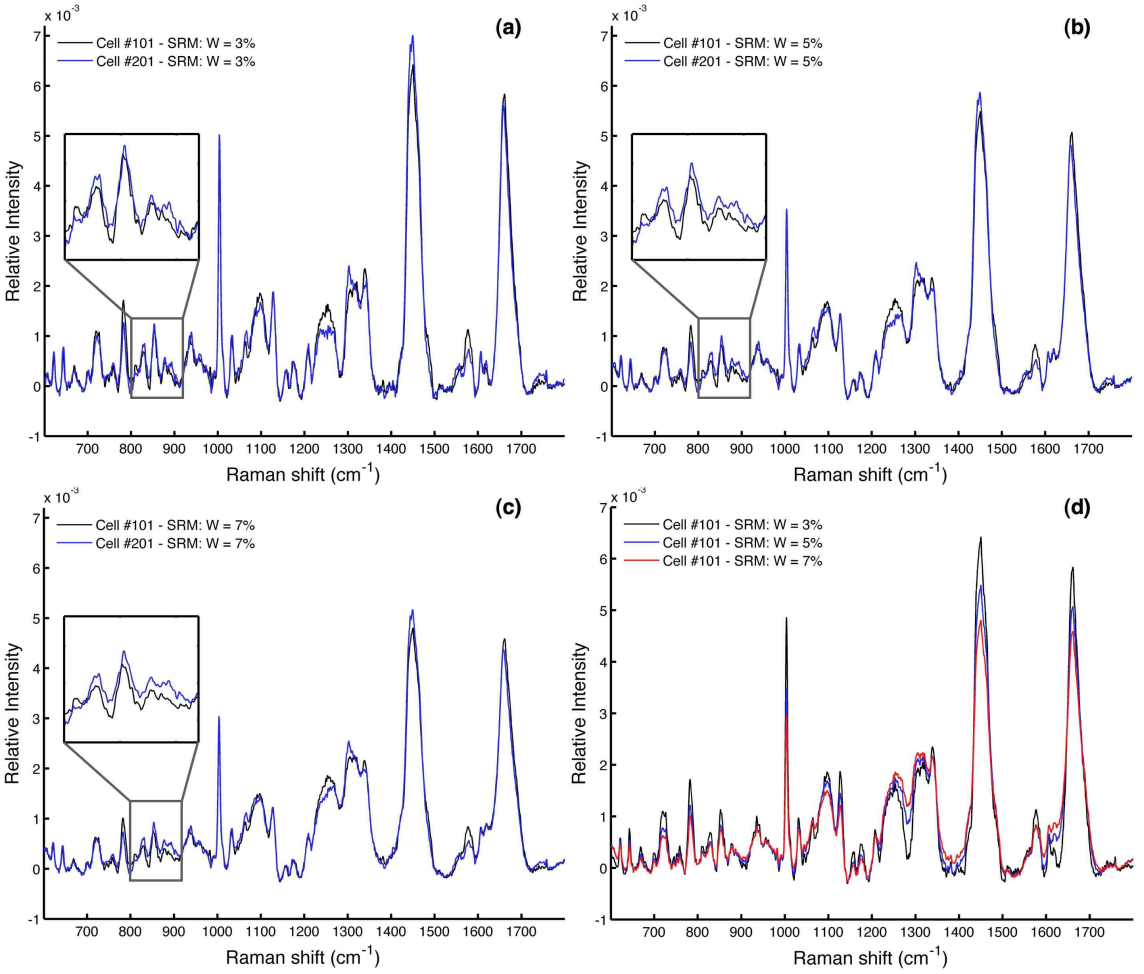


Figure 4.4: Effect of using the signal removal method (SRM) with single window sizes (W) of (a) 3%, (b) 5% and (c) 7% on two processed cell spectra, with a ROI (800-920 cm $^{-1}$) highlighted to show the variability in the subtracted BL level between the two spectra, in a region of inherently variable fluorescence between spectra. (d) BL subtracted spectra for a single cell using these window sizes.

the 1200 - 1800 cm^{-1} region during BL estimation than larger window sizes (figure 4.4). However, all three window sizes remove these Raman intensities in a fairly reproducible manner between spectra, as demonstrated by the consistent BL levels between the two spectra in the 1200 - 1800 cm^{-1} region (figures 4.4a to 4.4c). However, this reproducibility may not be consistent across all spectra in a large data set, which is investigated below.

4.3.2 Effect on PCA components and PCA scores

PCA components

The effect of using the SRM with single window sizes of 3%, 5% and 7% on the first two PCA components of the sample DU145 data set is shown in figure 4.5, with arrows indicating the wavenumbers (or regions) with observable differences in the features between the PCA components obtained with different window sizes. Two types of differences can be observed: (1) differences in the intensity (positive or negative) of features known to correspond to spectral variability in Raman signals, and (2) differences in the intensity of features (or regions) likely arising from variability in estimated BL levels. Each type is discussed in turn below.

The features at 853 cm^{-1} (PCA Component #1), and 877 cm^{-1}, 1340 cm^{-1} and 1421 cm^{-1} (PCA Component #2) are known to correspond to variability in molecular Raman signals (assigned in chapter 5). A reduction in the intensity of these features (*i.e.*, the intensity becomes closer to zero) indicates an undesirable loss in the sensitivity of the PCA algorithm to detect variability in the Raman signals at these wavenumbers. The features at 853 cm^{-1} (PCA Component #1) and 877 cm^{-1} (PCA Component #2) are most intense for W=3%, with losses in intensity observed for each increase in window size (figure 4.5). Since both these features are contained in the 800-920 cm^{-1} ROI shown in figure 4.4, it is likely that this loss of sensitivity to Raman spectral variability arises from the inability of the larger window BL estimates to conform to the valley points in this region. Conversely, the features at 1340 cm^{-1} and 1421 cm^{-1} (PCA Component #2) are most intense for W=7%, with the largest reduction in intensity occurring for W=3% (figure 4.5). Both these features occur at the shoulder regions of broad overlapping Raman signals (*e.g.*, figure 4.4), suggesting that a very conformal BL estimate (*i.e.*, W=3%) used in these regions removes Raman signals in an inconsistent manner, resulting in an undesirable loss of sensitivity to variability in Raman signals between spectra. Therefore, in terms of maintaining

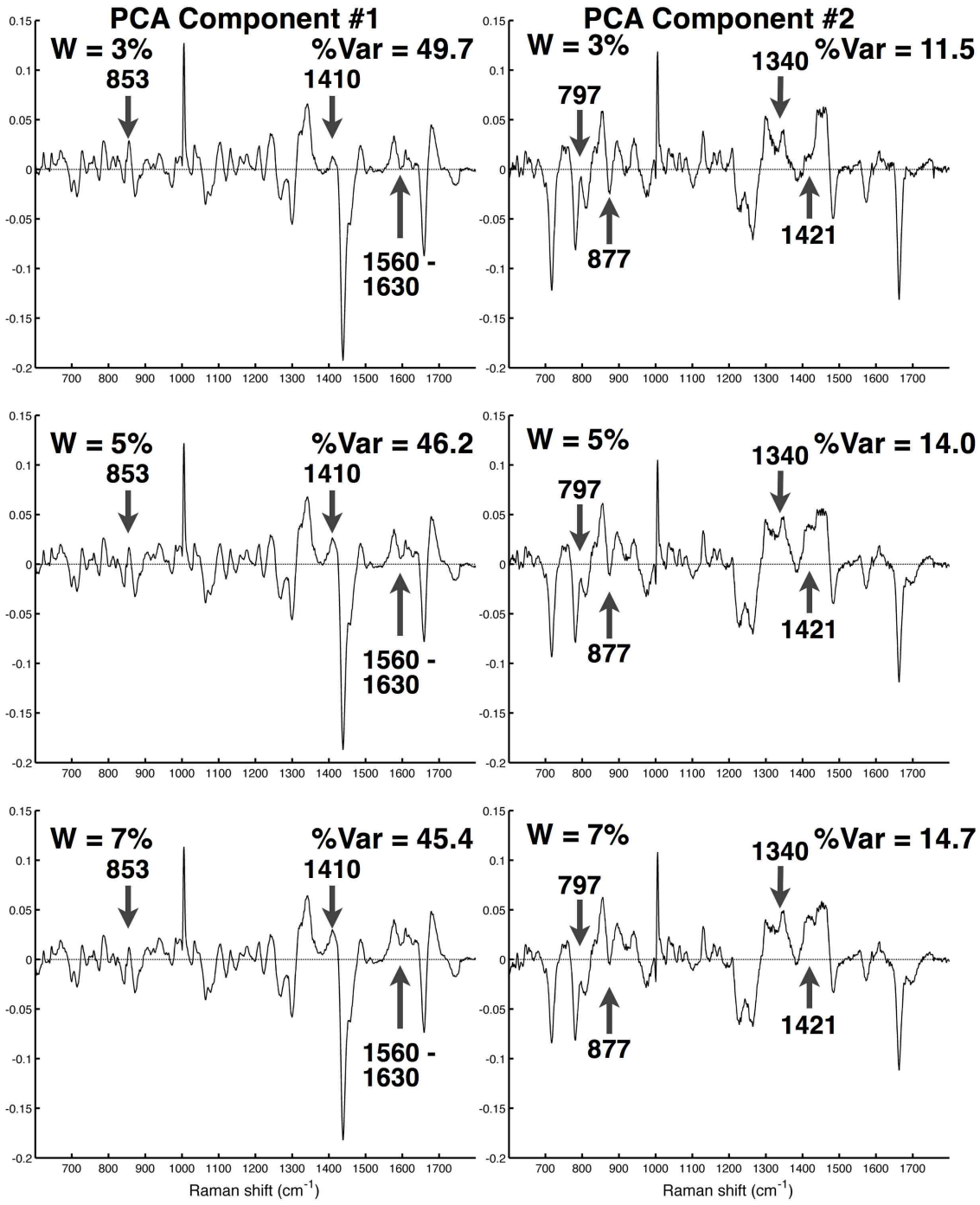


Figure 4.5: Effect of using the signal removal method (SRM) with single window sizes (W) of 3%, 5% and 7% on the first two PCA components of the sample DU145 data set. Arrows indicate the wavenumbers (or regions) with observable differences in features between the PCA components obtained using different window sizes.

sensitivity to variability in Raman signals, the SRM with W=3% performs best for the first half of the LWN spectral window, and the SRM with W=7% performs best for the second half of the LWN spectral window (although W=5% performs almost equivalently in this region).

The variable intensities of the features at 797 cm^{-1} (PCA Component #2) and 1410 cm^{-1} (PCA Component #1) likely arise from differences in estimated BL levels, as these regions are devoid of known Raman signals (see chapter 5). The region from 1560-1630 cm^{-1} (PCA Component #1) does contain Raman signals (*e.g.*, figure 4.4), but the variability between window sizes observed in figure 4.5 arises from variability in the BL level, not in the intensity of the Raman features above the BL. The SRM with W=3% performs best for minimizing the intensity of these features, for both PCA components. However, the loss in sensitivity to variability in known Raman signals when W=3% is used in the 1200-1800 cm^{-1} region (discussed above) may result in a loss of important molecular information, whereas a misidentification of the BL position is only problematic if the misrepresented BL level interferes with the ability to detect features arising from variability in Raman signals. For the PCA components shown in figure 4.5, the variability in the BL level at 1410 cm^{-1} and 1560-1630 cm^{-1} for W=5% and W=7% does not negatively affect the identification of the features arising from variability in Raman signals.

The PCA components produced for this sample data set in the remainder of this work (chapter 6), using W=3% for the first half (600-1190 cm^{-1}) and W=7% for the second half (1191-1800 cm^{-1}) of the LWN spectral window, are shown in figure 4.2 (for $X = 0.25$). Since the window sizes are set as percentages relative to the width of the spectral region, the SRM method used with a single window size produces the same BL estimate in a given region whether it is applied to the whole data set or half the data set. As such, the PCA components calculated in this work, using two different BL estimates for each half of the spectral window (figure 4.2, $X = 0.25$), are almost identical to the concatenation of the respective regions of the PCA components obtained when using single window sizes of W=3% and W=7%. Therefore, the benefits discussed above, for a certain window size applied to a certain region of the spectral window, are maintained when using different window sizes for different spectral regions. Furthermore, the PCA components obtained using the method of joining two different BL estimates (figure 4.2, $X = 0.25$) contain no visible features or artifacts at 1190 cm^{-1}, the position where the two BL estimates are joined (recall figure 3.9).

PCA scores

The effect of using the SRM with single window sizes of 3%, 5% and 7% on the PCA scores for the first two PCA components of the sample DU145 data set is shown in figure 4.6. Changes in the absolute values of the PCA scores cannot directly indicate meaningful differences resulting from using different window sizes, as the score values are calculated relative to the intensity of the original mean-subtracted spectra (recall equation 3.2). Since the overall intensity of the normalized spectra is larger for smaller window sizes (*e.g.*, figure 4.4), the values of the PCA scores, and the standard deviations of the PCA score distributions for each sample, are proportionately larger as well. This effect is observed in figure 4.6 by the decreased separations in the PCA score distributions between samples, and the decreased width of the distributions, with increased window size. However, this scaling does not directly affect the discriminatory capability of the PCA results, since the standard deviation of the PCA score distributions for each sample also decreases proportionately with the decreases in PCA scores. As such, the "relative distance" between the mean values of two PCA score distributions can be approximated by dividing the absolute distance between the mean values by the average standard deviation of all PCA score distributions in the data set, for a certain component. These relative distances can then be used, along with the p-values indicating the level of significance of the differences between PCA score distributions, to evaluate the effect of using these different window sizes for BL estimation on the discriminatory capability of the PCA results. Table 4.1 provides the absolute values of the relative distances between the mean values of adjacent PCA score distributions for the first two PCA components. The level of statistical significance for each shift between samples is indicated by the number of superscripted asterisks.

For the PCA scores for the first PCA component, there are no consistent changes in the relative distances between samples for different window sizes (table 4.1), and there is only one instance where the significance level drops (table 4.1, samples 6 & 7, W=7%). However, since the threshold levels for significance are set arbitrarily, and the p-values for W=5% and W=7% for these samples are 0.0003 and 0.0006, respectively, this isolated change in significance level does not indicate a loss in the discriminatory capability of the PCA results for the first PCA component.

For the PCA scores for the second PCA component, there is a consistent reduction in the relative distance between the mean values of adjacent PCA score distributions

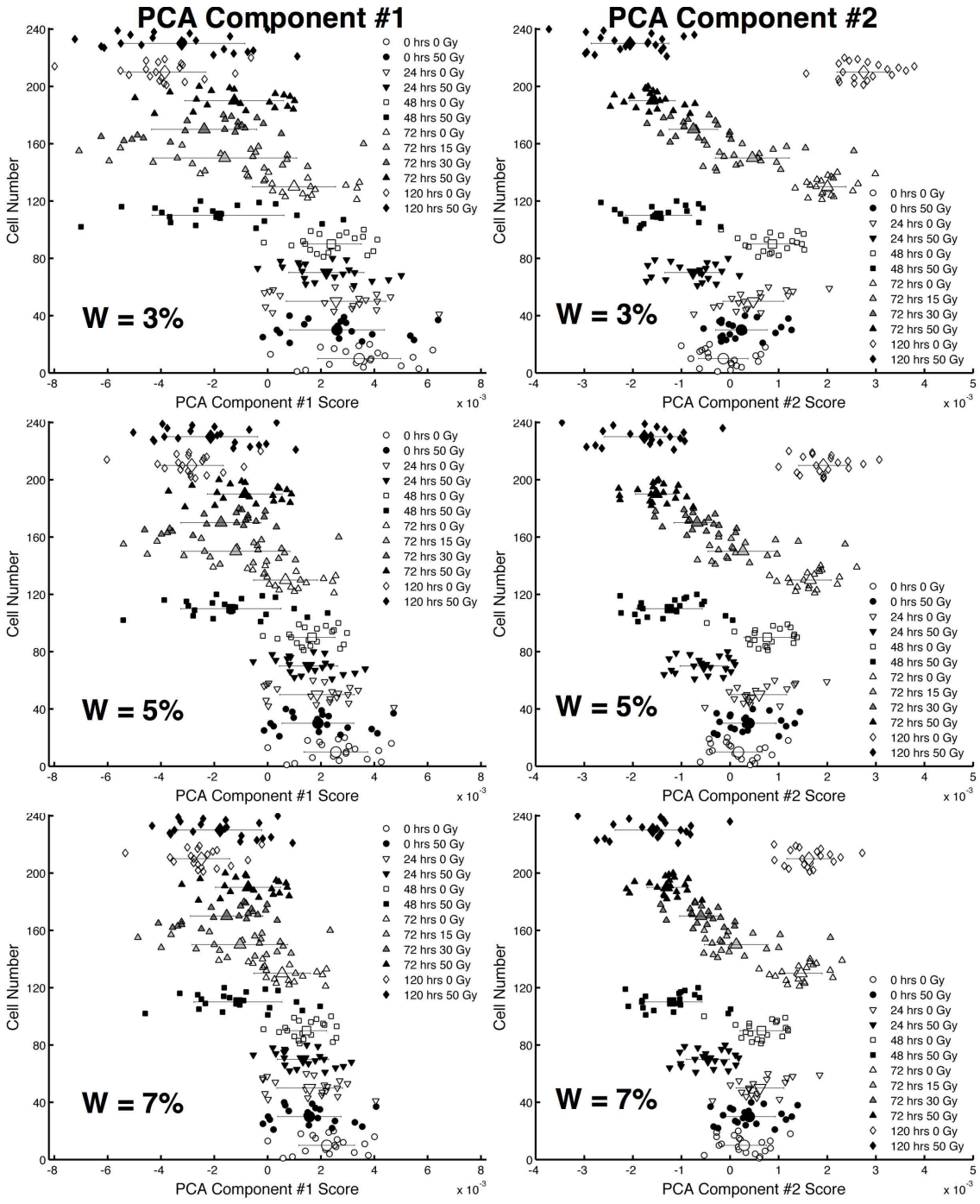


Figure 4.6: Effect of using the signal removal method (SRM) with single window sizes (W) of 3%, 5% and 7% on the PCA scores for the first two PCA components of the sample DU145 data set. Different markers categorize all 240 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

Table 4.1: Absolute values of the "relative distances" (see text) between the mean values of adjacent PCA score distributions (*i.e.*, samples) for the first two PCA components of the sample DU145 data set (figure 4.6), using the signal removal method (SRM) for baseline (BL) estimation with single window sizes (W) of 3%, 5% and 7%. Italicized values indicate data obtained using both W=3% and W=7% and the BL joining method applied for this work. The level of statistical significance for each shift between samples is indicated by the number of superscripted asterisks (none $= p > 0.05$, $^{*} = p \leq 0.05$, $^{**} = p \leq 0.005$, $^{***} = p \leq 0.0005$).

Samples	PCA Component #1				PCA Component #2			
W=	3%	5%	7%	*3&7%*	3%	5%	7%	*3&7%*
1 & 2	0.45	0.48	0.52	*0.47*	0.65*	0.40	0.20	*0.45*
2 & 3	0.02	0.02	0.01	*0.01*	0.43	0.37	0.43	*0.27*
3 & 4	0.20	0.23	0.20	*0.14*	2.17***	2.05***	2.09***	*1.88****
4 & 5	0.10	0.10	0.11	*0.13*	2.85***	2.36***	2.15***	*2.23****
5 & 6	2.29***	2.16***	2.07***	*2.10****	4.09***	3.63***	3.60***	*3.32****
6 & 7	1.54***	1.46***	1.34**	*1.36***	6.06***	5.23***	5.19***	*4.64****
7 & 8	1.40**	1.33**	1.27**	*1.34***	2.70***	2.50***	2.60***	*2.19****
8 & 9	0.42	0.39	0.40	*0.44*	2.12***	1.70***	1.42***	*1.73****
9 & 10	0.61	0.63	0.65	*0.62*	1.47***	1.52***	1.38***	*1.53****
10 & 11	1.41***	1.41***	1.41***	*1.37****	7.58***	6.20***	5.72***	*5.85****
11 & 12	0.34	0.50	0.57	*0.46*	8.38***	6.65***	6.27***	*6.13****

as the window size is increased, with only a few exceptions between W=5% and W=7% and only one exception between W=3% and W=5% (table 4.1). This may indicate a loss of the discriminatory capability of the PCA results for the second PCA component when less conformal BL estimates are used. However, there is only one instance of a change in the significance level for different window sizes, occurring between samples 1 & 2 (table 4.1) where the p-value drops from 0.51 to 0.19 to 0.03 as the window size is decreased from W=7% to W=5% to W=3%. The reason for this effect is unclear, although it is interesting to note that for all the further results presented in this work for similarly conducted radiation experiments (including other cell lines), there are no statistically significant shifts in the PCA score distributions between sample 1 (0 hrs 0 Gy) and sample 2 (0 hrs 50 Gy). As such, this effect may be reflective of some bias induced into the PCA score distributions as the result of using a very conformal BL estimation method in regions containing broad Raman features. As indicated by the data obtained when using both W=3% and W=7% and the BL joining method (table 4.1), samples 1 & 2 are not statistically different using the analysis methods applied for this work. Furthermore, it can be seen in figure 4.6 that the separation between samples 1 & 2 for W=3% is in the opposite direction as the subsequent larger separations between unirradiated and irradiated samples harvested for RS analysis at later time points (24-120 hrs). Since the source of variability causing the differences in the PCA score distributions arises from the same signature of spectral variability for each sample (described by the second PCA component, figure 4.5), it is likely that the effect noted here is indeed the result of some bias arising from the use of a very conformal BL estimation method in regions containing broad Raman features.

4.4 Spectral normalization

Spectral normalization is performed in this work using the total area under the BL subtracted spectrum [70, 166, 167]. Some authors have addressed this issue of intensity variability by normalizing to the area under the peak centred at 1450 cm^{-1} [59, 81, 83], claiming that the area of this peak arises solely from CH bonds common to all biomolecules, and is therefore proportional to the total amount of biological material within the sampling volume. However, for the sample DU145 data set used here (and most subsequent data sets), the most significantly varying Raman signal between spectra, as indicated by the most intense feature in the first PCA components,

occurs at 1438 cm^{-1} (strongest negative feature in PCA component #1, figures 4.2 and 4.5). The Raman signal at 1438 cm^{-1} arises predominantly from lipids (assigned in later chapters), and therefore does not scale with the amount of total biological material in the sampling volume if the lipid content changes relative to other types of biomolecules (*i.e.*, proteins, nucleic acids, *etc.*). Therefore, the area of the 1450 cm^{-1} peak may not be suitable for normalization, due to the presence of overlapping spectral features, when investigating the molecular sources of spectral variability between single-cell spectra. The effect of normalizing to the area of the 1450 cm^{-1} peak as compared to normalizing to the total area under the spectrum, after BL subtraction, is presented below. For this analysis, the area under the 1450 cm^{-1} peak was determined by integrating the intensity of the BL subtracted spectrum from 1440 to 1460 cm^{-1}.

4.4.1 Effect on processed spectra

The differences between the relative intensities of two spectra when normalized by the total area ("TA") or by the area of the 1450 cm^{-1} peak ("A1450") are shown in figure 4.7. The ROI highlights the 1450 cm^{-1} peak region for each normalization method, and the difference spectrum is plotted for each case to highlight the effect of the different normalizations on the resulting variability between the two spectra. Since the absolute intensity of each normalized spectrum depends on the normalization value, the Y-axis limits of each plot (figures 4.7a and 4.7b) are scaled relative to the maximum intensity of the cell #201 spectrum for each case.

The ROI highlighted in figure 4.7 clearly indicates the difference in relative intensity between the two spectra, in the 1450 cm^{-1} peak region, for each normalization method. This immediately indicates that the two normalization methods do not equivalently represent the total amount of biological material within the sampling volume, since the overall intensity of the cell #201 spectrum is reduced relative to the cell #101 spectrum when normalizing to the area under the 1450 cm^{-1} peak (figure 4.7). The effect on the Raman variability between spectra can be seen by comparing the two difference spectra. The Raman features exhibiting the largest changes between the two methods are indicated by the arrows in figure 4.7. In the A1450 difference spectrum (figure 4.7b), relative to the TA difference spectrum (figure 4.7a), the strong negative feature at 1430-1470 cm^{-1} is not observed, the negative feature at 1300 cm^{-1} is reduced in intensity, and the positive features at 784, 1094-

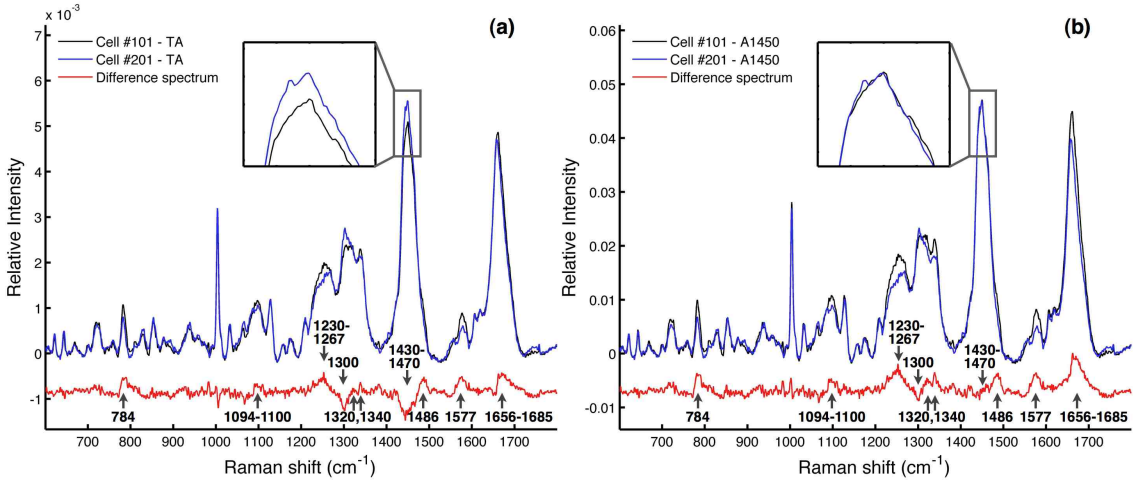


Figure 4.7: Effect of normalizing to (a) the total area (TA) or (b) the area under the 1450 cm^{-1} peak (A1450) on two processed DU145 cell spectra, with the 1450 cm^{-1} peak region highlighted for each normalization method. The difference spectrum is plotted for each case, with wavenumbers provided for the most significantly altered features.

1100, 1230-1267, 1320, 1340, 1486, 1577 and 1656-1685 cm^{-1} all exhibit an increase in intensity. As such, the interpretation of the changes in Raman signals between spectra is affected by the normalization method.

It is important to note that the features of reduced intensity in the A1450 difference spectrum have significant contributions from lipids (*i.e.*, 1300 cm^{-1} and 1438 cm^{-1}), whereas the features of increased intensity arise from proteins and nucleic acids (assigned in chapter 5). Therefore, it is likely that using the 1450 cm^{-1} peak for normalization reduces the sensitivity to changes in Raman signals from lipids between spectra, as the 1450 cm^{-1} peak area contains varying amounts of Raman signals arising from lipids at 1438 cm^{-1}. The variability observed in the A1450 difference spectra therefore reflects variations arising from proteins and nucleic acids relative to lipid content, with a loss in sensitivity to changes in Raman signals arising from lipids throughout the LWN spectral window. Furthermore, observed changes in protein and nucleic acid signals may not reflect an accurate change in the concentration of these biomolecules within the sampling volume, relative to other spectra, as the changes will be biased by the varying contributions to the normalization value arising from lipids.

4.4.2 Effect on PCA components and PCA scores

PCA components

The effect of normalizing by the total area (TA) or by the area of the 1450 cm^{-1} peak (A1450) on the first two PCA components of the sample DU145 data set is shown in figure 4.8, with arrows indicating the wavenumbers (or regions) with the largest differences in the features between the PCA components obtained with each method.

For the first PCA components obtained using different normalization methods (figure 4.8), the most significant differences observed for the A1450 component, relative to the TA component, are the reductions in the intensities of the negative features at 700, 719, 877, 1065, 1080, 1267, 1300, 1438, 1450, 1460 and 1656 cm^{-1}. In the A1450 component, some of these features are reduced to approximately zero intensity, and some are newly represented as positive features. As shown in chapter 5, all of these features arise from variability in Raman signals originating from a variety of lipid molecular groups. Therefore, as discussed above during analysis of the two normalized spectra, using the 1450 cm^{-1} peak area for normalization reduces the sensitivity to changes in Raman signals from lipids throughout the LWN spectral window.

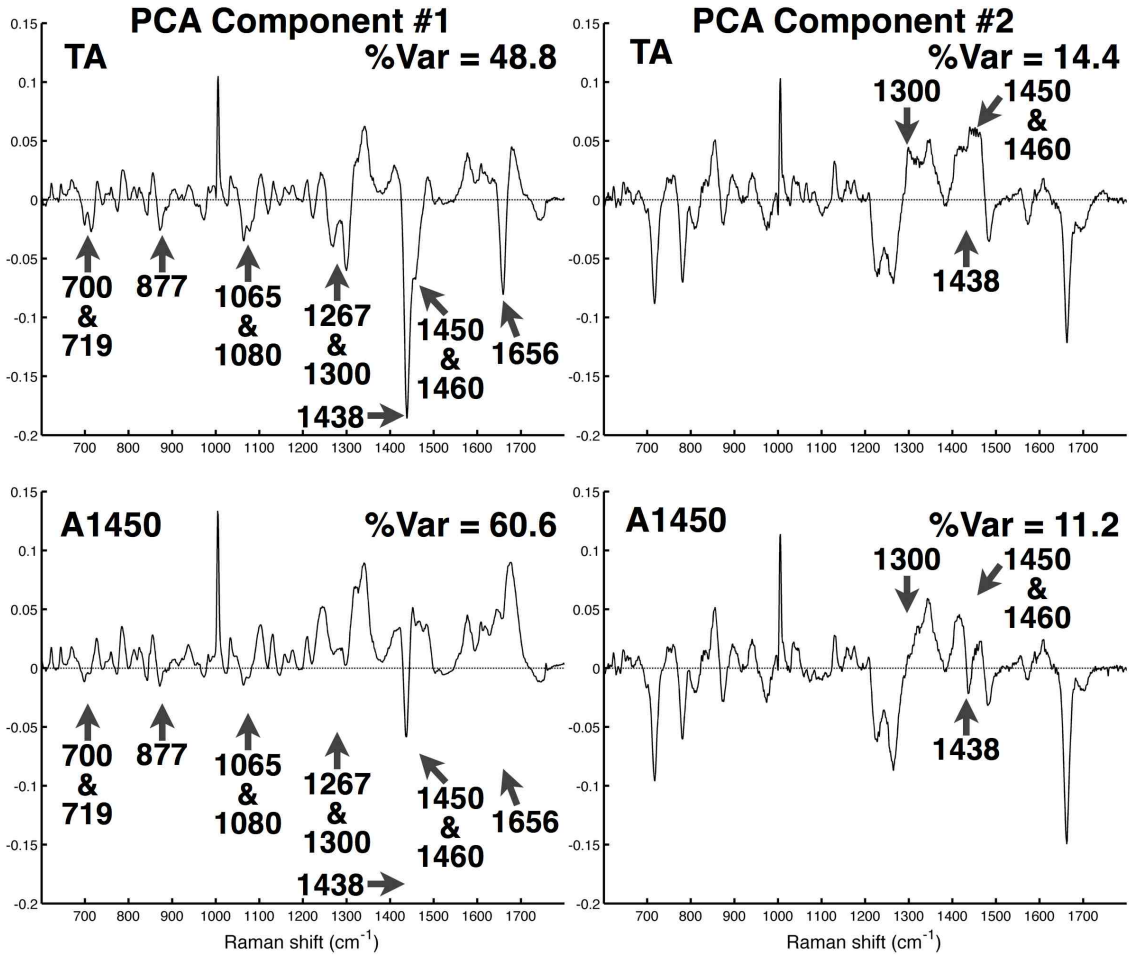


Figure 4.8: Effect of normalizing by the total area (TA) or by the area of the 1450 cm^{-1} peak (A1450) on the first two PCA components of the sample DU145 data set. Arrows indicate the wavenumbers (or regions) with the largest differences in features between the PCA components obtained with each method.

For the second PCA components obtained using different normalization methods (figure 4.8), the most significant differences observed for the A1450 component, relative to the TA component, are the reductions in the intensities at 1300, 1438, 1450 and 1460 cm^{-1}. The reductions in the 1440-1460 cm^{-1} region are expected as this is the region used for normalization, therefore eliminating much of the variability in this region prior to the PCA calculation. However, the reduction in the intensity of the features at 1300 and 1438 cm^{-1} further indicates a loss of sensitivity to changes in Raman signals arising from lipids when the 1450 cm^{-1} peak area is used for normalization.

PCA scores

The effect of normalizing by the total area (TA) or by the area of the 1450 cm^{-1} peak (A1450) on the PCA scores for the first two PCA components is shown in figure 4.9. As was the case for the different BL estimation parameters investigated in section 4.3.2, the normalization value affects the magnitude of the PCA scores since the PCA score values are calculated relative to the intensity of the original mean-subtracted spectra. As such, the X-axis limits in figure 4.9 are scaled relative to the range of the PCA score values, allowing visual comparisons of the trends in the data.

There are no apparent changes in the trends of the PCA score distributions for each PCA component between normalization methods, although there are observable shifts in the positions of individual distributions, most notably for the first PCA component (figure 4.9). For the second PCA component, the width of the distributions for each sample appears to be reduced for the A1450 method, relative to the apparent separation between distributions. These effects on the discriminatory capability of the PCA results can be quantified by the "relative distance" and p-value approach, as used above for the analysis of the PCA scores obtained using different BL estimation parameters (table 4.1). The absolute values of the relative distances between the mean values of adjacent PCA score distributions for the first two PCA components, obtained when normalizing by the total area (TA) and by the area of the 1450 cm^{-1} peak (A1450), are provided in table 4.2. As before, the level of statistical significance for each shift between samples is indicated by the number of superscripted asterisks.

For the PCA scores for the first PCA component, there is no consistent trend of changes in the relative distances between samples when using a different normalization method (table 4.2), although there are several instances of differences in the level

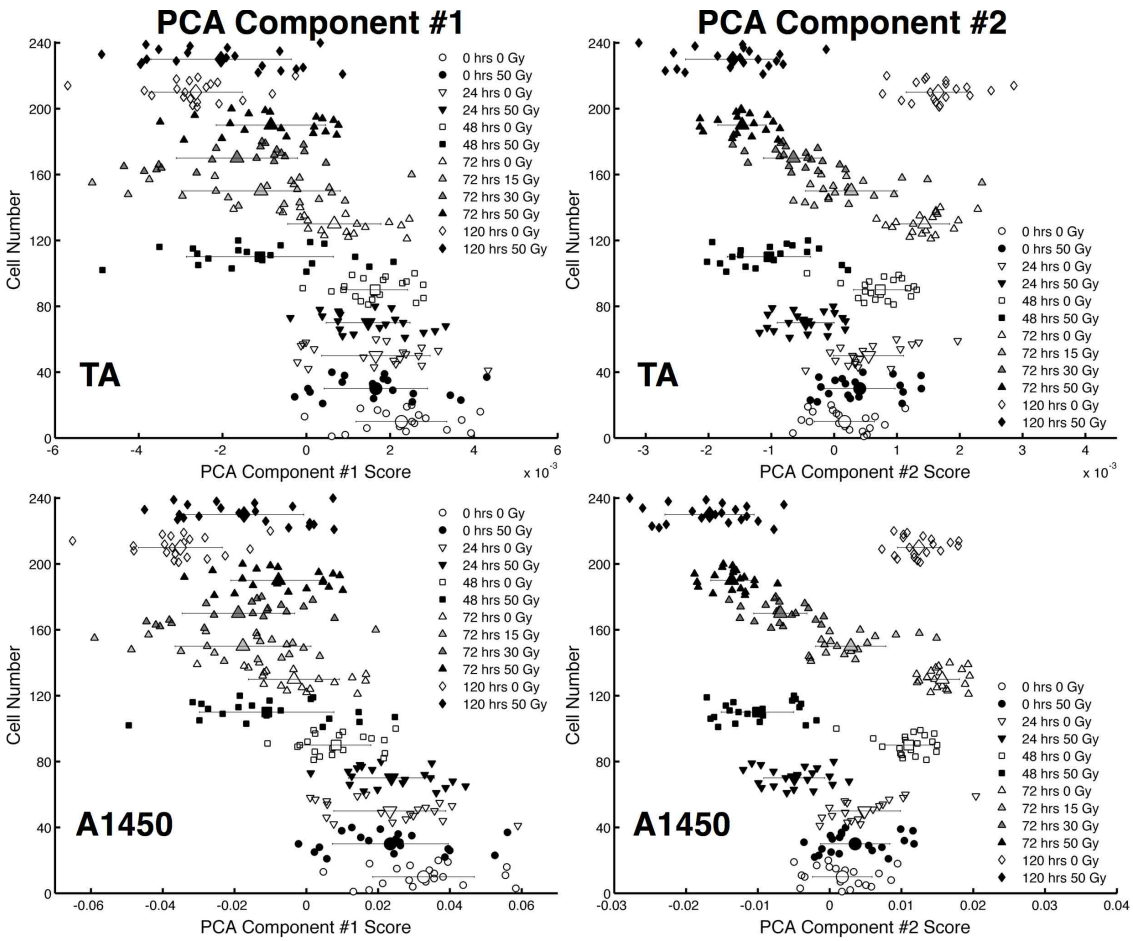


Figure 4.9: Effect of normalizing by the total area (TA) or by the area of the 1450 cm^{-1} peak (A1450) on the PCA scores for the first two PCA components of the sample DU145 data set. Different markers categorize all 240 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

Table 4.2: Absolute values of the "relative distances" (see text) between the mean values of adjacent PCA score distributions (*i.e.*, samples) for the first two PCA components of the sample DU145 data set (figure 4.6), when normalizing by the total area ("TA") or by the area of the 1450 cm^{-1} peak ("A1450"). The level of statistical significance for each shift between samples is indicated by the number of superscripted asterisks (none = $p > 0.05$, $^{*} = p \leq 0.05$, $^{**} = p \leq 0.005$, $^{***} = p \leq 0.0005$).

Samples	PCA Component #1		PCA Component #2	
	TA	A1450	TA	A1450
1 & 2	0.47	0.63	0.45	0.43
2 & 3	0.01	0.01	0.27	0.32
3 & 4	0.14	0.03	1.88^{***}	2.36^{***}
4 & 5	0.13	1.05^{***}	2.23^{***}	3.84^{***}
5 & 6	2.10^{***}	1.32^{***}	3.32^{***}	5.08^{***}
6 & 7	1.36^{**}	0.52	4.64^{***}	6.20^{***}
7 & 8	1.34^{**}	0.97^{*}	2.19^{***}	3.07^{***}
8 & 9	0.44	0.09	1.73^{***}	2.36^{***}
9 & 10	0.62	0.76^{*}	1.53^{***}	1.64^{***}
10 & 11	1.37^{***}	1.87^{***}	5.85^{***}	6.29^{***}
11 & 12	0.46	1.21^{***}	6.13^{***}	7.03^{***}

of significance between adjacent samples (table 4.2). This result is not surprising, since many features in the first PCA component are affected by the normalization method used (figure 4.8). As such, the PCA scores for the first PCA component are discriminating based on a different set of parameters (*i.e.*, the different features in the components) for each normalization method. The general trend of the PCA score distributions between samples (figure 4.9) is preserved since there are still many positive and negative features that are common to both of the first PCA components obtained for each normalization method (figure 4.8).

For the PCA scores for the second PCA component, the relative distance between adjacent samples is consistently larger for the A1450 method relative to the TA method, with exceptions between samples 1 & 2 and 2 & 3 where there is no significant difference between the PCA score distributions for either normalization method. The reason for this increase in the relative distance between PCA score distributions is unclear, although it may indicate an improvement in the discriminatory capability of the PCA results obtained when normalizing to a specific peak area, rather than to the total area under the BL subtracted spectrum. However, this potential improvement in discrimination between samples coincides with a loss in sensitivity to variability in Raman signals originating from molecules that contribute signals to the peak used for normalization. Therefore, in terms of retaining sensitivity to the molecular origins of Raman spectral variability, normalizing to the total area under the BL subtracted spectrum is likely a preferable approach.

4.5 Discussion

For the RS studies performed in this work, the most important outcomes of the data analysis are (1) the detection of molecular sources of spectral variability as represented by the features in the PCA components, and (2) the discriminatory capability of the PCA results as represented by the separations between PCA score distributions for a given PCA component. The PCA score distributions for each sample indicate the presence of molecular changes between samples, and the corresponding PCA components indicate the molecular groups that are changing between samples. These outcomes ultimately determine the merit of using RS for the biochemical analysis of single human tumour cells, and form the basis for the assignment of molecular origins to any radiation response observed with RS. As such, the data processing steps and parameters used in this work were chosen to minimize any detrimental effects on these

two important outcomes of the data analysis. The variations in the data processing parameters for each processing step presented in this chapter have varying effects with respect to these data outcomes, as discussed below.

A high level of spectral smoothing, with the TPMEM smoothing parameter $X = 1.25$, is found to reduce sensitivity to variability in the intensity of sharp Raman features, relative to no smoothing, or moderate smoothing with $X = 0.25$. Moderate smoothing with $X = 0.25$ is found to maintain sensitivity to spectral variability for all Raman features, while appreciably reducing the level of random noise in the data set and possibly increasing sensitivity to subtle sources of spectral variability. Analysis of the PCA score distributions demonstrates that the tested levels of spectral smoothing have no measurable effect on the discriminatory capability of the PCA results for the first two PCA components.

The analysis of the first two PCA components obtained using the SRM for BL estimation with different window sizes demonstrates that in terms of maintaining sensitivity to variability in Raman signals, the SRM with W=3% performs best for the first half of the LWN spectral window and the SRM with W=7% performs best for the second half of the LWN spectral window. Furthermore, the analysis of the PCA score distributions for the first two PCA components indicates that a bias may be induced into the relative separations between PCA score distributions as the result of using a very conformal BL estimation method in regions containing broad Raman features, possibly biasing the discriminatory capability of the PCA results. Within the parameters tested here, the optimal BL estimation method is to concatenate two different BL estimations for each half of the LWN spectral window, thus maintaing the benefits of each BL estimation for each spectral region and avoiding any possible biases caused by inconsistent removal of Raman signal intensity.

The analysis of the two different methods of normalization indicates that the normalization method may affect the interpretation of the molecular sources of spectral variability between spectra. In this analysis, such changes are observed in both the normalized spectra and the first two PCA components. Specifically, it is demonstrated that normalizing to the area of a single Raman peak reduces sensitivity to variability in the Raman signals that contribute to the peak chosen for normalization. Conversely, normalizing to the total area under the BL subtracted spectrum, which is proportional to the total Raman intensity collected from the sampling volume, does not restrict sensitivity to any particular source of spectral variability, although it may slightly reduce the discriminatory capability of the PCA results relative to

single-peak normalization methods. In this work, however, a confident assignment of the molecular sources of spectral variability is paramount. As such, the total area under the BL subtracted spectrum is used for normalization in this work.

4.6 Conclusion

A representative set of raw data is used to demonstrate the effect of varying certain processing parameters, related to spectral smoothing, baseline estimation and spectral normalization, on the final outcomes and interpretations of the analysis. The data processing steps and parameters used in this work are chosen to minimize any detrimental effects on (1) the detection of molecular sources of spectral variability and (2) the discriminatory capability of the PCA results. The analysis presented in this chapter demonstrates the robustness and validity of the developed techniques for the application of PCA to large Raman data sets ($\sim$200 spectra) obtained from single cells.

Chapter 5

Results & Discussion II: Variability in Raman spectra of single human tumour cells

5.1 Introduction

This chapter presents work published by the Society of Applied Spectroscopy (Frederick, MD, USA) in the journal *Applied Spectroscopy*, entitled "Variability in Raman spectra of single human tumor cells cultured *in vitro*: correlation with cell cycle and culture confluency" [168], and is reproduced here with permission from the Society of Applied Spectroscopy.

In a given population of cells there are biochemical differences between individual cells due to a number of reasons, such as cells existing at different points in their cell cycle (in the case of an asynchronously growing cell culture), or cells growing in cultures of different confluence (in the case of multiple cell cultures). Several studies, mentioned in section 1.3.2, have used RS to analyze biochemical differences arising within a population of a single type of cell. These studies characterized spectral features obtained from dead cells [59], apoptotic cells [81], necrotic cells [82], and non-proliferative cells [83], and demonstrated successful discrimination between G0/G1 cells and S or G2/M cells [84].

In the studies mentioned above, natural biochemical differences exist between individual cells in a sample. In all these previous studies (excepting the cell cycle discrimination study [84]), spectral variability arising from inherent biochemical dif-

ferences was either averaged during spectral acquisition or post-processing, or was not a relevant or necessary consideration for the purpose of the study. However, the spectral differences between cell populations may be very subtle, and it is important to thoroughly investigate the existing sources of spectral variability within a given population. To date, there have been no systematic studies on the inherent sources of RS spectral variability that may arise due to biochemical differences between single cells.

This chapter presents an investigation of the capability of RS to detect inherent sources of spectral variability within a human tumour cell line (DU145) cultured *in vitro*. PCA is used to observe differences in Raman spectra that correlate with cells existing in different phases of the cell cycle, as well as differences that correlate with the confluency of the cell culture at the time of Raman analysis. Furthermore, the biochemical changes detected by RS that correlate with cell cycle progression are consistent with known biochemical changes within cells. It is shown that the variability from cell cycle and culture confluency comprises almost all of the inherent variability in a multi-culture data set, and is primarily explained by the first two PCA components obtained from the data.

The results of this chapter are presented in two studies. In the first study, the inherent variability between cultures of varying confluence is investigated by acquiring Raman spectra of cells collected from asynchronously growing cell cultures harvested 1-8 days after sub-culturing. In the second study, the variability due to cell cycle in sub-confluent cultures is directly examined by acquiring spectra from cells collected from cultures synchronized at specific points in the cell cycle. Flow cytometry is used to monitor the cell cycle distribution and viability of all cultures at the time of Raman analysis. RS spectra are obtained from adherent cells that have been resuspended and centrifuged into a pellet, from which individual cells are selected using a high power focusing objective and 785 nm laser excitation (recall section 3.2.2). This technique provides a very high quality Raman spectrum of a single cell, while eliminating any spectral variability (caused by inconsistent cell structure and varying local cell density) that arises when cells are grown and analyzed directly in a monolayer [169]. Both low-wavenumber (LWN), 600-1800 cm^{-1}, and high-wavenumber (HWN), 2700-3100 cm^{-1}, spectral windows are used in order to determine if information can be attained equivalently from either window, as some authors have found to be the case for certain applications [170].

5.2 Materials & Methods

5.2.1 Cell preparation

Asynchronous cell cultures

A single T-75 flask of DU145 cells was grown to ∼90% confluency and sub-cultured equally into 8 identical T-75 flasks. The first flask was harvested for RS analysis 24 hours after sub-culturing, and each remaining flask was harvested every 24 hours thereafter. Each time a flask was harvested, the cell cycle distribution and viability of the culture was measured as described previously (section 3.1.5).

Cell culture confluency is defined as the percentage of the surface area of the culture flask covered by cells. Confluency estimates of each of the asynchronous cell cultures were obtained prior to harvesting by acquiring low magnification images of five different regions of the cell culture. Each image was imported into Matlab and the fraction of the image covered by cells was calculated. Since the confluency is never consistent throughout the entire surface area of the flask, the averaged confluency from the five regions is used as an estimate of the overall confluency of the culture.

Synchronized cell cultures

Following established protocols [171, 172], 2 mM of thymidine (Sigma-Aldrich, Oakville, ON, Canada) was used to inhibit DNA replication and arrest cells at the end of G1 phase, before the onset of S phase, and 100 ng/mL of nocodazole (Sigma-Aldrich) was used to prevent formation of the mitotic spindle and arrest cells at the end of G2 phase, prior to M phase. In the second study presented in this chapter, cell cultures were treated with thymidine and nocodazole to obtain cultures synchronized at four points during the cell cycle: (1) at the G1/S boundary (with double-thymidine [171] treatment), (2) during S phase (with double-thymidine treatment and re-incubation with drug-free media for 3 hours), (3) at the G2/M boundary (with thymidine treatment for 24 hours, re-incubation with drug-free media for 3 hours, nocodazole treatment for 12 hours, and mitotic shake-off) and (4) during early G1 phase (with mitotic shake-off from a G2/M arrested culture and re-incubation of detached cells with fresh media for 5 hours).

5.2.2 Raman spectroscopy and data processing

Single-cell RS acquisition and data processing was performed as previously described (sections 3.2.2 and 3.3), using both the LWN (600-1800 cm^{-1}) and the HWN (2700-3100 cm^{-1}) spectral windows. PCA calculations were performed separately on the LWN and HWN window data sets to facilitate an independent comparison and corroboration of results obtained from each window.

In both studies presented in this chapter, spectral variability arising from the quartz substrate was easily identifiable in a single PCA component obtained from the LWN window; the quartz component was therefore removed and the PCA calculations were repeated on the filtered set of spectra, as described in section 3.3.5.

5.3 Results

5.3.1 Single DU145 cell spectrum

In the LWN spectral window the Raman spectrum of a single DU145 cell (figure 5.1a) contains multiple contributions from proteins, lipids and nucleic acids. Spectral features of proteins arise from aromatic amino acids (phenylalanine, tryptophan and tyrosine), amide groups of secondary protein structures (α-helices, β-sheets and random coils), and the stretching or deformation of carbon atoms bonded with nitrogen, hydrogen or other carbon atoms. Nucleic acid features include contributions from individual RNA and DNA bases (adenine, thymine, guanine, cytosine and uracil), as well as from the sugar-phosphate backbone of DNA. A number of different lipid features are also detectable throughout the spectral window.

In the HWN window, the spectrum (figure 5.1b) is a superposition of broad features dominated by the stretching of various lipid and protein CH_2 and CH_3 groups. There are also weak contributions from =CH stretching in lipids and from aromatic groups in both nucleic and amino acids. A detailed listing of the molecular assignments [59, 75, 92, 118–121] for all spectral features observed in this work is provided in table 5.1.

Liquid water yields broad and weak Raman signals in the spectral ranges 100-900 cm^{-1}, 1400-1800 cm^{-1} and 3000-3700 cm^{-1} [173]. In this work, signals from water likely contribute to the spectral baseline (recall figures 3.9 and 3.10), along with fluorescence from biological material and the quartz substrate.

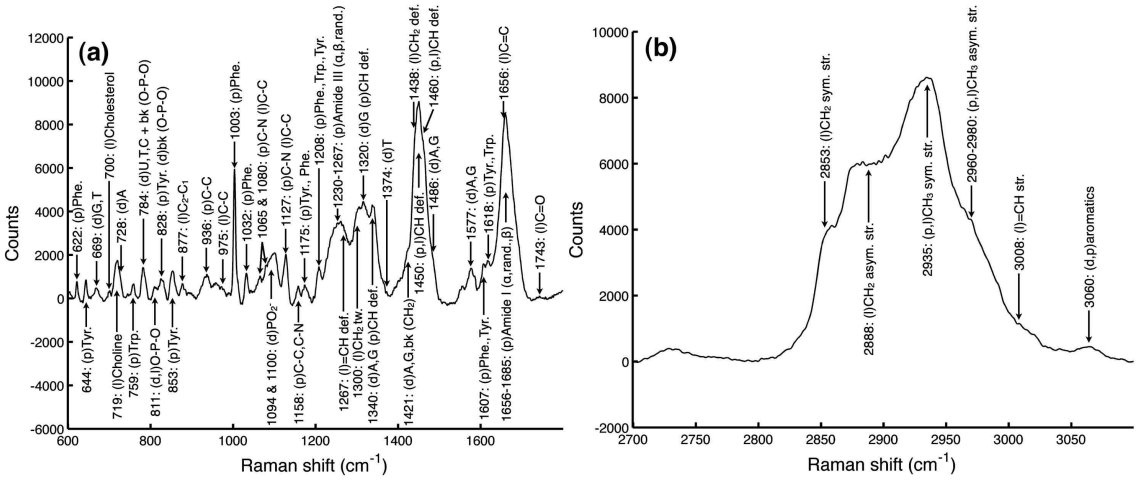


Figure 5.1: Raman spectra of a single DU145 cell for the (a) LWN and (b) HWN spectral windows; the Raman shift and molecular origin of identifiable features are provided [59, 75, 92, 118–121]. Abbreviations - p: protein, l: lipid, d: DNA/RNA, A: adenine, T: thymine, G: guanine, C: cytosine, U: uracil, Phe: phenylalanine, Tyr: tyrosine, Trp: tryptophan, bk: backbone, def: deformation, tw: twist, sym: symmetric, asym: asymmetric, str: stretch.

Table 5.1: Molecular assignments for spectra of DU145 cells. Superscript numbers indicate references used for particular assignment.

Raman Shift ($\mathbf{cm^{-1}}$)	**Molecular Assignment**		
	DNA/RNA	**Proteins**	**Lipids**
622		C-C tw Phe[118, 119]	
644		C-C tw Tyr[118, 119]	
669	G,T[118, 119]		
700			Cholesterol[119]
719			Choline[118, 119, 121]
728	A[118–120]		
759		Trp ring br[118, 119]	
784	U, C, T ring br[118–120] O-P-O bk[118, 120]		
811	O-P-O RNA[118]		O-P-O[121]
828	O-P-O asym[118, 120]	Tyr ring br[92, 118–120]	
853		Tyr ring br[92, 118–120]	
877			Acyl C_2-C_1[118, 121]
936		C-C sym bk, α[75, 92, 118–120]	
975			Head C-C[118, 121]
1003		Phe sym ring br[75, 92, 118–120]	
1032		C-H Phe[75, 118, 120]	
1065		C-N[118, 120]	Chain C-C[75, 92, 118, 119, 121]
1080		C-N[75, 118, 120]	Chain C-C[75, 92, 118, 121]
1094	PO_2^- bk, B[75, 118–120]	C-N[75, 120]	Chain C-C[92, 118, 121]
1100	PO_2^- bk, A[92, 119]		
1127		C-N[75, 118, 120]	Chain C-C[75, 92, 118, 121]
1158		C-C, C-N[118]	
1175		C-H Tyr, Phe[75, 118, 120]	
1208		C-C_6H_6 Phe, Trp, Tyr[75, 92, 118–120]	
1230		Amide III rand coil[59]	
1246		Amide III β[59, 92]	
1255		Amide III β, α[59, 75, 92, 119, 120]	
1267		Amide III α[59, 92]	=CH def[59, 121]
1300			CH_2 tw[75, 92, 118, 119, 121]
1320	G[75, 118]	CH def[59, 118]	
1340	A, G[75, 118–120]	CH def[59, 118]	
1374	T[75, 119, 120]		
1421	A, G[75, 120], CH_2 bk[119]		
1438			CH_2 def[75, 92, 119, 121]
1450		CH def[59, 75, 92, 118–120]	CH def[59, 118]
1460		CH def[59, 118]	CH def[59, 118]
1486	A, G[119, 120]		
1577	A, G[75, 118–120]		
1607		C=C Phe, Tyr[118]	
1618		C=C Tyr, Trp[118]	
1656		Amide I α[59, 92]	C=C[59, 75, 92, 119, 121]
1660		Amide I[75, 118–120]	
1669		Amide I rand coil[92]	
1685		Amide I β[92]	
1743			C=O[92, 118, 119, 121]
2853			CH_2 sym[75, 92, 119–121]
2888			CH_2 asym[75, 92, 119–121]
2935		CH_3 sym[75, 92]	CH_3 sym[75, 119, 121]
2960–80		CH_3 asym[75, 92]	CH_3 asym[75, 92, 119, 121]
3008			=CH[92, 121]
3060	aromatics[92]	aromatics[92]	

Abbreviations - A: adenine, U: uracil, C: cytosine, T: thymine, G: guanine, Trp: tryptophan, Tyr: tyrosine, Phe: phenylalanine, br: breathing, bk: backbone, def: deformation, tw: twist, sym: symmetric, asym: asymmetric.

5.3.2 Study #1: Asynchronous cell cultures

Cell cycle, confluency and viability

The cell cycle distributions and culture confluencies for the 8 samples in this study (figure 5.2) are typical for asynchronous cells growing to confluency in culture. From 24-72 hours after sub-culturing the distribution among the three phases is fairly constant at ∼50% G1, ∼20% G2 and ∼30% S. Between 72 and 96 hours there is an increase in the G1 phase fraction and a decrease in both the G2 and S phase fractions. After 96 hours the G2 fraction remains relatively constant, whereas the S fraction decreases, and the G1 fraction increases, until about 168 hours. One element that is not measurable with this method of cell cycle analysis is the fraction of cells in 'G0' phase, a state of cellular quiescence. Cellular quiescence is only achievable during G1 phase, usually soon after cell division; therefore, G0 cells are indistinguishable from G1 cells by the flow cytometry methods used in this study.

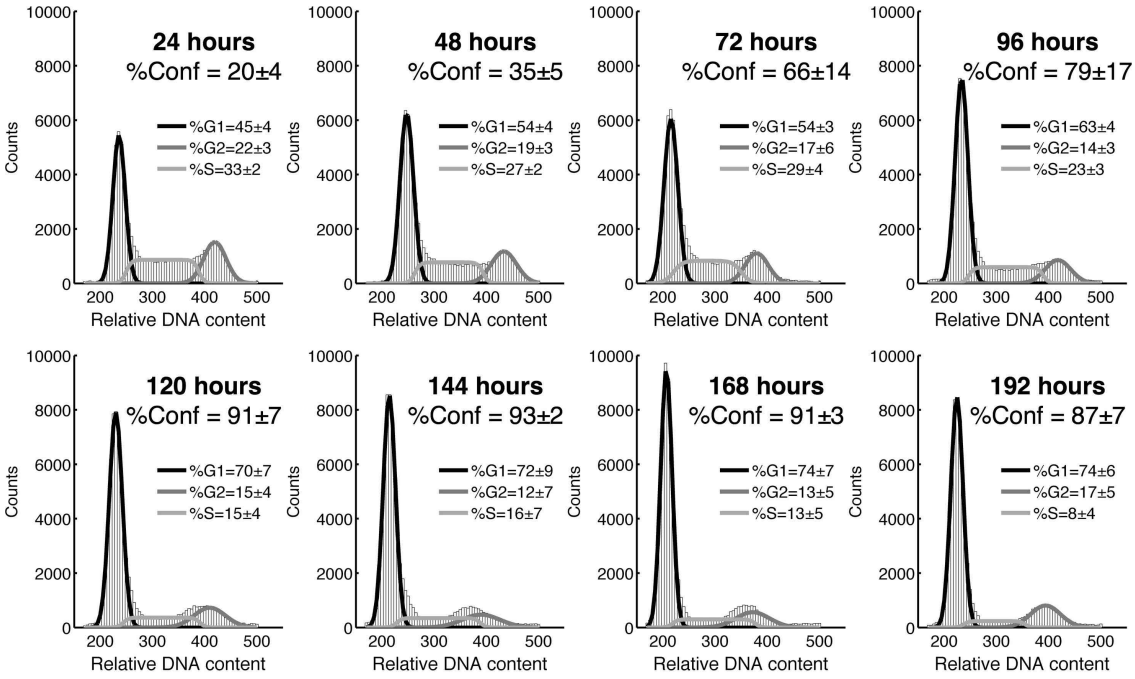


Figure 5.2: Flow cytometry analysis of cell cycle distributions for the asynchronous cell cultures. Time indicates the incubation time of the culture after sub-culturing. Culture confluency (%Conf) and cell cycle phase fractions were calculated as described in sections 5.2.1 and 3.1.5.

The viability of the harvested cells was determined with flow cytometry prior to Raman analysis. Dead cells usually detach from the growth substrate, and are

subsequently rinsed off and discarded during the harvesting procedure. However, a small percentage of dead cells always remains in a harvested culture. For this study, viability tests proved that all of the first 7 samples (24-168 hours after sub-culturing) had a viability of >98% (*i.e.*, less than 2% dead cells), and the 8th and final sample (192 hours after sub-culturing) had a viability of >95%.

First PCA component

For the 160 cell spectra collected in this study, the first PCA components (figure 5.3) represent the most significant source of spectral variability in each data set (52.6% of the total variance for the LWN window, 88.6% for the HWN window). By comparison with the known Raman shifts (figure 5.1, table 5.1), the features in the PCA components for both the LWN and HWN window are identifiable as arising from variability in the Raman intensity of peaks in the original data set; therefore, molecular origins can be assigned to the features in the components. The PCA components consist of both positive and negative features; any spectrum assigned a higher (*i.e.*, more positive) PCA score for a given component contains a proportionately higher amount of the positive features, and a lower amount of the negative features, from that component.

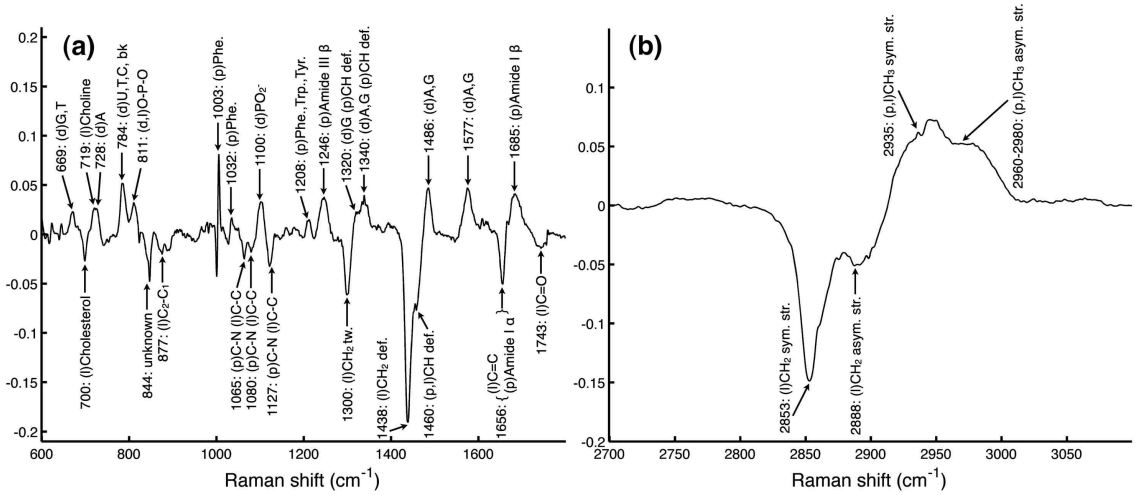


Figure 5.3: First PCA components from the asynchronous cell cultures study: (a) LWN window (52.6% of total variance), (b) HWN window (88.6% of total variance). The Raman shift and molecular origin of identifiable features are provided [59, 75, 92, 118–121]. Abbreviations - as in figure 5.1.

The negative features in the first PCA component for the LWN window (figure

5.3a) are dominated by lipid contributions from cholesterol, CH_2 twisting, CH_2 and CH deformation, and C-C, C=C and C=O stretching. The C-C features at 1065, 1080 and 1127 cm^{-1}, the CH feature at 1460 cm^{-1}, and the C=C feature at 1656 cm^{-1} have the same Raman shifts as the protein contributions from C-N stretching, CH deformation, and α-helix amide groups, respectively. As such, it is impossible to determine whether or not this variability is solely lipid or protein in nature, or some combination of both; however, as the rest of the negative features in the component are uniquely lipid in origin, it follows that the negative nature of these features arises in part from lipids as well. The molecular origin of the negative feature at 844 cm^{-1} is unknown. The positive features in the LWN component are almost exclusively nucleic acid and protein in origin, with the exception of a weak positive contribution from choline at 719 cm^{-1}. Nucleic acid features arise from DNA and RNA bases, and from the DNA backbone. Protein features arise from aromatic amino acids (phenylalanine, tryptophan and tyrosine) and from β-sheet amide groups. Interestingly, it is known that the aromatic amino acids are most likely to be found in a β-sheet conformation, and less likely to be found in an α-helix or coiled structure [174]. As such, an increase in signal from the aromatic amino acids should correlate with an increase in signal from β-sheet amide groups, which is observed to be the case here, provided that the amino acids are used in the synthesis of secondary protein structures. For the HWN window (figure 5.3b), the positive features can be assigned to the symmetric and asymmetric stretching of CH_3 groups in both proteins and lipids. The negative features, however, arise from the symmetric and asymmetric stretching of CH_2 groups in lipids alone. To summarize, the negative features in both the LWN and the HWN window are primarily due to lipids, whereas the positive features in both the LWN and the HWN window are primarily due to nucleic acids and proteins (in particular from amino acids and β-sheet amides for the LWN window).

The PCA scores (figure 5.4) determine how much of the variability explained by the first components (figure 5.3) is expressed in each of the 160 cell spectra in each data set. Note that positive scores are correlated with increased nucleic acid and protein content, and negative scores are correlated with increased lipid content. The scores for both windows show the same overall trend; between 24 and 48 hours after sub-culturing there is a significant increase ($p = 0.005$ (LWN), $p = 0.003$ (HWN)) in nucleic acid and protein content relative to lipid content, followed by a steady decrease in the average nucleic acid and protein content relative to the average lipid content from 48 to 192 hours after sub-culturing. Furthermore, the relative positions

of the individual cell scores are consistent between the LWN and HWN windows. For example, cells #30 and #39 (figure 5.4) have, respectively, the highest and lowest scores from the 48 hours LWN window sample set, and the same two cells have, respectively, the highest and lowest scores from the corresponding HWN window sample set. It is worth emphasizing that the intra-sample variability in the PCA scores for a given PCA component arises from the same source of spectral variability as the inter-sample variability, and simply reflects the intrinsic biochemical heterogeneity of each cell culture.

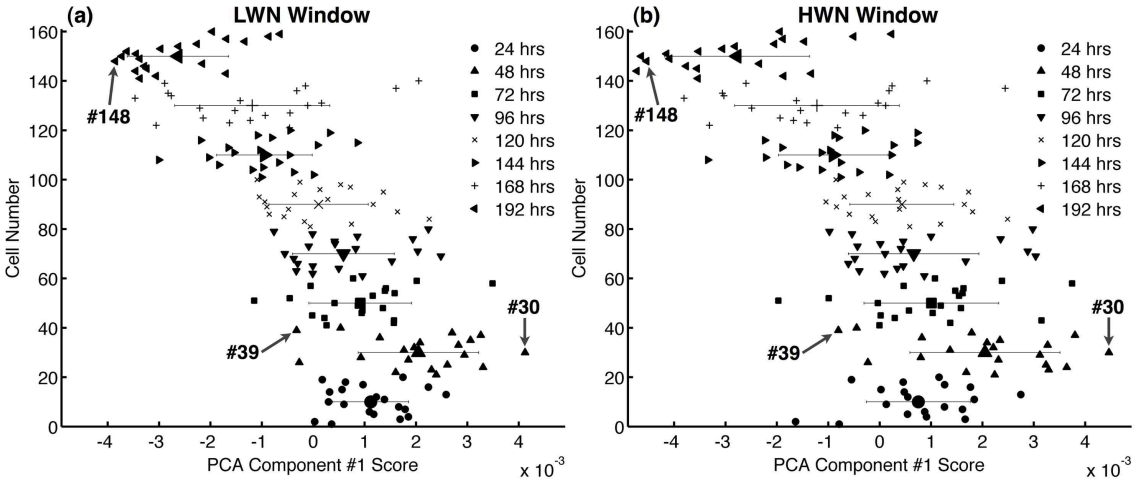


Figure 5.4: PCA scores for the first components from the asynchronous cell cultures study, for the (a) LWN and (b) HWN window. Scores for all 160 cells are grouped by time of harvest after sub-culturing. The average score and standard deviation is shown for each sample for visualization of the trends in the data. The Raman spectra of cell #30 and cell #148 are shown in figure 5.5.

To show that the variability observed with PCA is directly observable in the original data, the Raman and difference spectra for two cells (cells #30 and #148) that have a large separation in their PCA scores (figure 5.4) are shown in figure 5.5, along with the PCA components for comparison. All of the major features in the components are directly observable in the corresponding difference spectrum for each spectral window, without any rescaling of the difference spectra.

Second PCA component

The second PCA component for the LWN window (figure 5.6a) explains 10.1% of the total variance, and the corresponding component for the HWN window (figure 5.6b) explains 1.7% of the total variance. Assigning a molecular origin to the features in

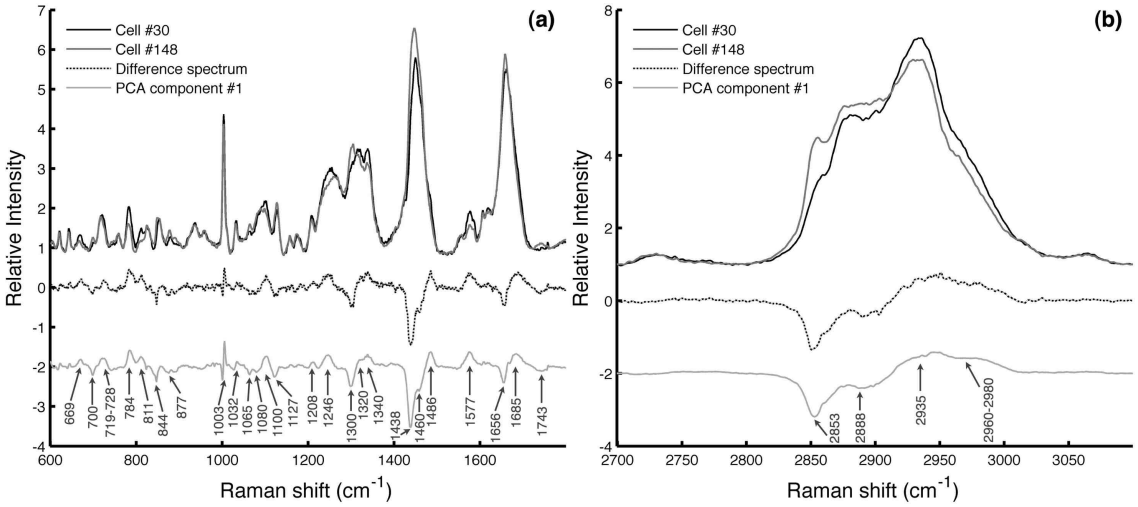


Figure 5.5: Raman and difference spectra for two cells (#30 and #148) with a large difference in PCA score (figure 5.4) for the first PCA component. The first PCA components have been offset and rescaled for comparison with the unscaled difference spectra. Wavenumbers are provided for any known features in the components (figure 5.3) that are also observable in the difference spectra.

the second components is more difficult than for the first components, especially for the HWN window where the percent variance explained is very low and there is a small number of known molecules contributing to the HWN spectra (figure 5.1b, table 5.1). The features in the HWN window that correspond with known wavenumbers are the symmetric and antisymmetric stretching of protein and lipid CH_3 groups at 2935 and 2975 cm^{-1}, respectively, and the antisymmetric stretching of lipid CH_2 groups at 2888 cm^{-1} (figure 5.6b). The accuracy of these assignments is uncertain, especially since contributions from protein and lipid CH_3 groups appear in both the positive and negative features of the component. However, for the LWN window (figure 5.6a) most of the major features can be assigned with confidence. The negative features include a strong contribution from choline, as well as contributions from O-P-O stretching in lipids and RNA, the nucleic acid bases A and G, and a combined contribution from lipid =CH deformation and α-helix amide groups. The sharp negative feature at 1660 cm^{-1} arises from amide groups as well, but whether it arises from a certain protein conformation, or from amide groups in general, is unknown. The positive features arise from amino acids, amide groups in β-sheet and random coil conformation, and a combined contribution from the nucleic acid bases A and G and CH deformation in proteins. Since the nucleic acids A and G contribute uniquely to the negative features of the component, this combined positive contribution likely arises from CH deformation in proteins alone. The origin of the positive feature at 1120 cm^{-1} is unknown.

Despite the uncertainty in the molecular origins of the features in the second PCA components (especially for the HWN window), the PCA scores for both windows (figures 5.7a and 5.7b) still show the same general trend; from 24 to $\sim$120 hours after sub-culturing there is an overall increase in the average scores, and after $\sim$120 hours the average scores appear to remain relatively constant until decreasing slightly between 168 and 192 hours. However, the relative positions of the individual cell scores between the LWN and HWN windows are not consistent. Therefore, the similar trends between the two windows may not be the result of the same biomolecular changes occurring within the cells.

To determine if the variability described by the second PCA components is directly observable in the original data (as it is for the first PCA components (figure 5.5)), the Raman and difference spectra for two cells (cells #137 and #19 for the LWN window, and cells #114 and #31 for the HWN window) that have a large separation in their PCA scores (figure 5.7) are shown in figure 5.8, along with the PCA components

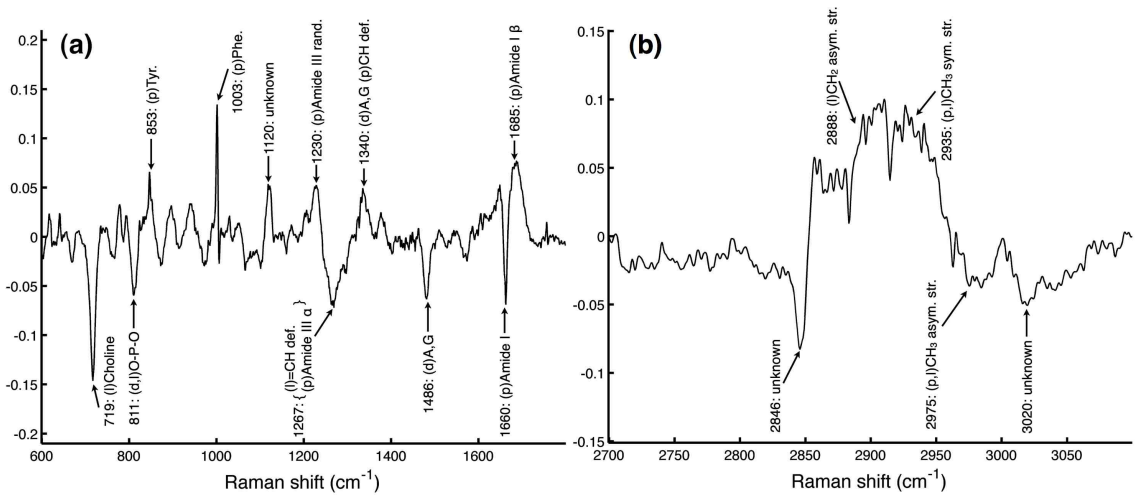


Figure 5.6: Second PCA components from the asynchronous cell cultures study: (a) LWN window (10.1% of total variance), (b) HWN window (1.7% of total variance). The Raman shift and molecular origin of identifiable features are provided [59, 75, 92, 118–121]. Abbreviations - as in figure 5.1.

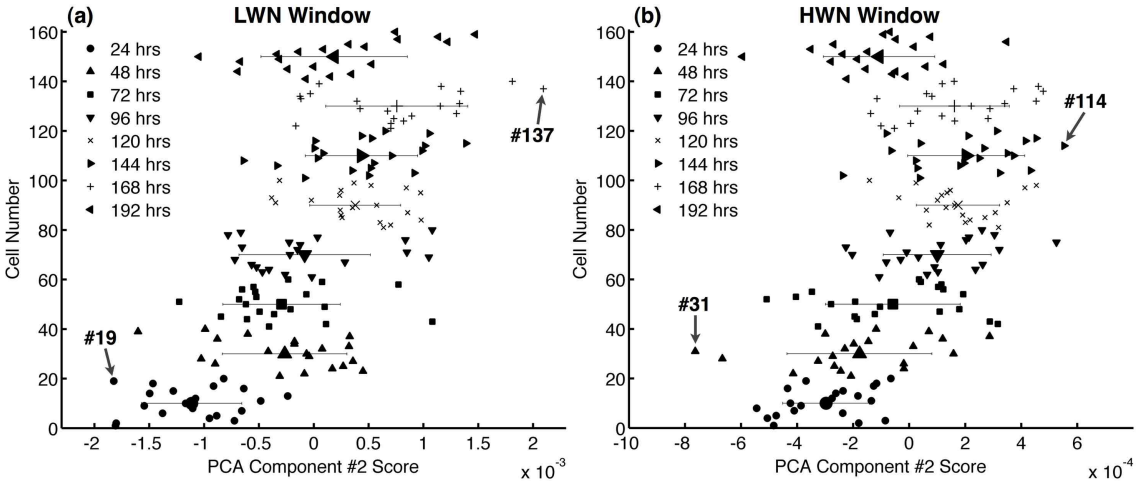


Figure 5.7: PCA scores for the second components from the asynchronous cell cultures study, for the (a) LWN and (b) HWN window. Scores for all 160 cells are grouped by time of harvest after sub-culturing. The average score and standard deviation is shown for each sample for visualization of the trends in the data. The Raman spectra of cells indicated by arrows are shown in figure 5.8.

for comparison. For the LWN window, all of the major features in the component are observable in the LWN difference spectrum. However, the features in the HWN component are not observable in the HWN difference spectrum.

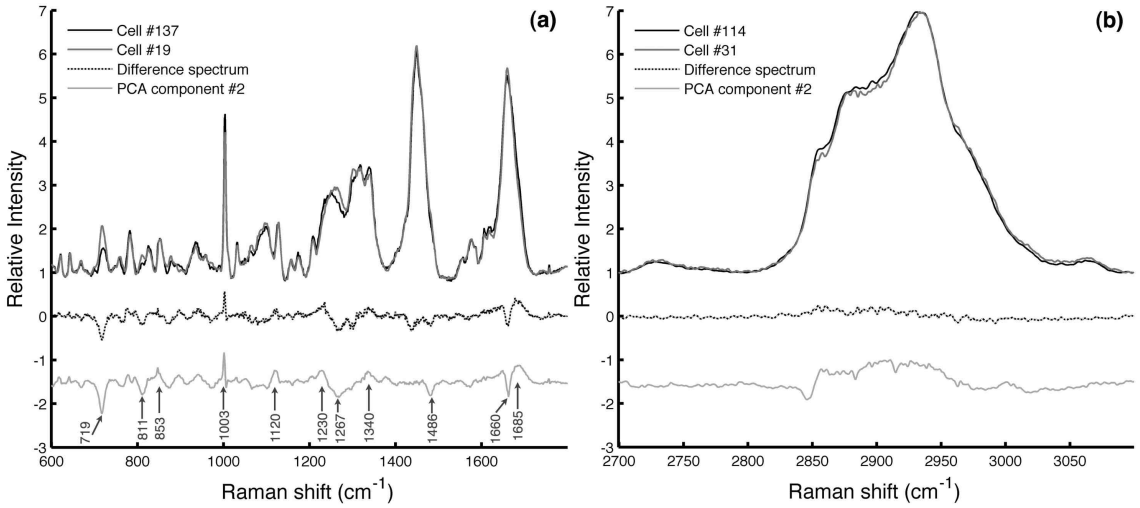


Figure 5.8: Raman and difference spectra for two cells (#137 and #19 for the LWN window, and #114 and #31 for the HWN window) that have a large difference in PCA score (figure 5.7) for the second PCA component. The second PCA components have been offset and rescaled for comparison with the unscaled difference spectra. Wavenumbers are provided for any known features in the components (figure 5.6) that are also observable in the difference spectra.

Other PCA components

The third PCA component for the LWN window (figure 5.9a) explains 7.6% of the total variance, and is dominated by a sharp derivative-like feature centred at the wavenumber of the sharp phenylalanine ring breathing peak at 1003 cm^{-1}. This feature in the PCA component indicates variability arising from a shift in the calibration of the Raman system over time. The trend displayed by the score plots for this component (figure 5.9b) correlates well with a known drift in the Raman calibration over the 8 day sample collection period, which was monitored by measuring peak shifts of the 520 cm^{-1} feature of an instrument-based silicon sample before and after the daily Raman collection. Furthermore, there are trends observed within the PCA score distributions for the 48, 72 and 196 hrs samples (figure 5.9b), which match the known direction of drift in the Raman calibration measured before and after the 2 hour collection period for each sample.

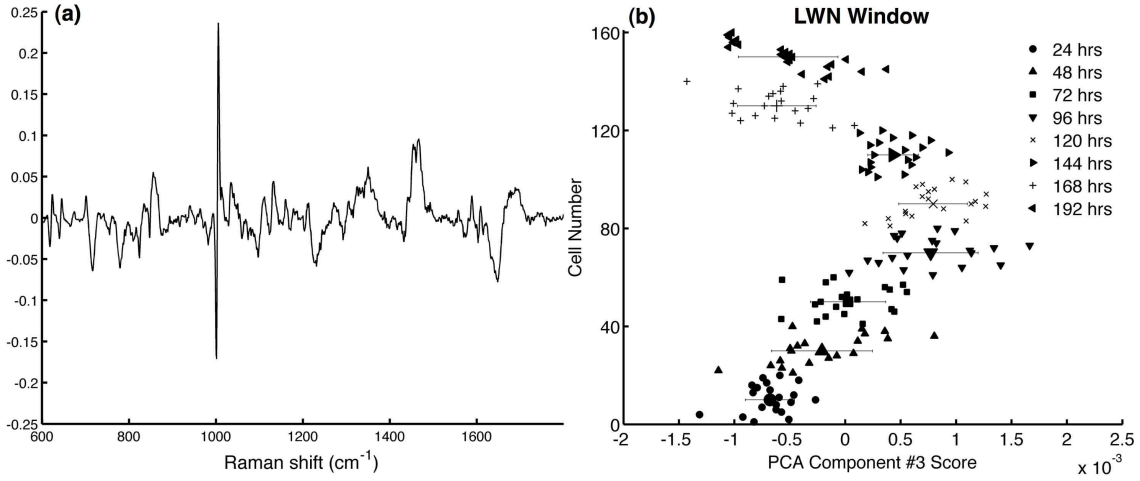


Figure 5.9: (a) Third PCA component (7.6% of total variance) and (b) PCA scores, for the LWN spectral window, from the asynchronous cell cultures study. Scores for all 160 cells are grouped by time of harvest after sub-culturing. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

During daily collections, it was verified that the initial calibration of the system was within 0.5 cm^{-1} of the calibration performed on the first day of collection. By inspection of various pairs of spectra with large differences in their scores for the third PCA component, it was found that the maximum shift in the position of the phenylalanine peak at 1003 cm^{-1} was less than one pixel (1 pixel $\approx$ 0.9 cm^{-1}, at 1003 cm^{-1}) for all 160 spectra collected during the 8 days of data collection. For spectra with the most outlying scores for the third PCA component, corrections to the shift were attempted using linear interpolation and were successful in reducing by a few percent the total amount of variance explained by the third PCA component. However, due to the sharpness of many peaks in the LWN spectrum there are always slight shifts measured in the peak positions due to experimental limitations, which translate into some amount of variability brought out in the PCA analysis.

In this study, the third PCA component is the last component for the LWN window that displays any measurable trend in the score plots; furthermore, each of the remaining 156 components explain less that 3% of the total variance, and likely have little to no biological significance. The same can be said for the remaining 157 PCA components for the HWN window, each of which explain less than 1% of the total variance. The remaining PCA components account for residual variance arising from random sources of spectral variability.

5.3.3 Study #2: Synchronized cell cultures

Cell cycle synchronization

Cell cultures were synchronized at four different points in the cell cycle: at the G1/S boundary, at 3 hours into S phase, at the G2/M boundary, and at early G1 phase. For the first culture, ~83% of the cells were successfully arrested either in late G1 or early S phase (figure 5.10a, 'G1/S'). Three hours after release from an identical G1/S arrest, ~19% of the second culture remained in G1 phase, whereas ~64% of the culture was measured to be in S phase (figure 5.10b, 'G1/S +3 hrs'). For the third culture a distinct G1 peak was not observed after G2/M synchronization. Therefore, the combined fraction of cells in G1 or S phase was estimated to be ~26%, with at least 74% of the cells successfully arrested at the G2/M boundary (figure 5.10c, 'G2/M'). Five hours after harvesting and re-incubating cells from an identical G2/M arrest, ~21% of the fourth culture was determined to be left in G2 phase, while ~75% of the culture was found in G1 phase (figure 5.10d, 'G2/M +5 hrs'). Since the fourth

culture was seeded with cells that were primarily at the G2/M boundary, the G1 cells in the fourth culture must be less than 4-5 hours into G1 phase.

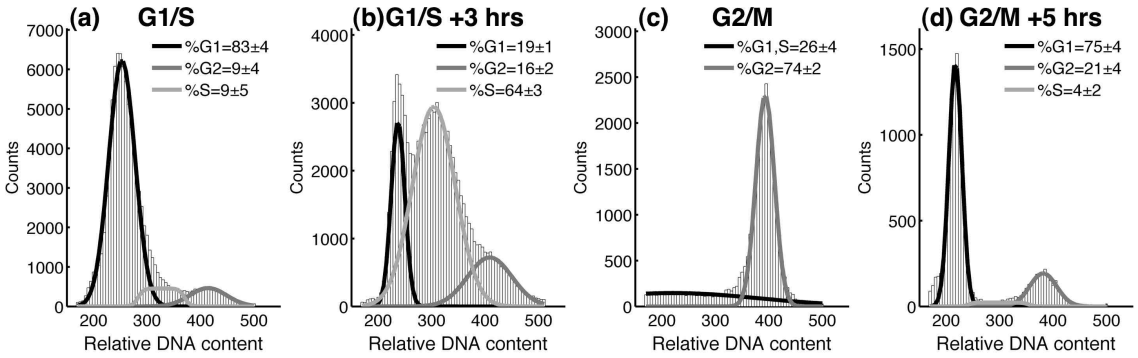


Figure 5.10: Flow cytometry analysis of cell cycle distributions for the synchronized cell cultures. Synchronization was performed using thymidine and nocodazole as described in section 5.2.1.

First PCA component

The first PCA component for the LWN window (figure 5.11a) explains 51.6% of the total variance and is very similar to the corresponding component from the asynchronous cell cultures study (figure 5.3a, 52.6% of the total variance). As in the previous study, the negative features in the component are dominated by lipid contributions from cholesterol, CH_2 twisting, CH_2 and CH deformation, and C-C, C=C and C=O stretching, with an additional negative contribution from choline, which previously contributed as a weak positive feature in the asynchronous study. There is also a new negative feature at 1267 cm^{-1}, a combined contribution from lipid =CH deformation and α-helix amide groups; this feature correlates with the existing negative combined contribution from lipid C=C stretching and α-helix amides at 1656 cm^{-1}. The previously observed negative features at 844 and 1127 cm^{-1} are not observed here. The positive features in the LWN component, as in the previous study, are exclusively nucleic acid and protein in origin, with contributions from DNA and RNA bases, the DNA backbone, aromatic amino acids and β-sheet amide groups. In this study there are additional positive contributions from tyrosine at 853 cm^{-1}, thymine at 1374 cm^{-1}, and random coil amide groups at 1230 cm^{-1}. The previously observed positive feature at 811 cm^{-1} is not observed here. The first PCA component for the HWN window (figure 5.11b) explains 86.6% of the total variance and is nearly identical to the corresponding component from the asynchronous cell cultures study

(figure 5.3b, 88.6% of the total variance). As before, the positive features arise from the symmetric and asymmetric stretching of CH_3 groups in both proteins and lipids, whereas the negative features arise from the symmetric and asymmetric stretching of CH_2 groups in lipids alone.

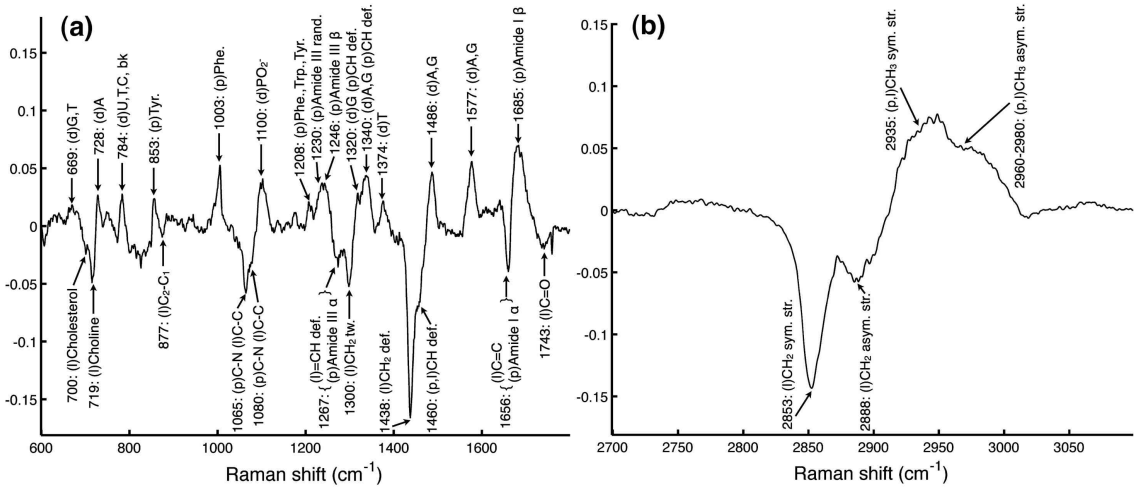


Figure 5.11: First PCA components from the synchronized cell cultures study: (a) LWN window (51.6% of total variance), (b) HWN window (86.6% of total variance). The Raman shift and molecular origin of identifiable features are provided [59, 75, 92, 118–121]. Abbreviations - as in figure 5.1.

The PCA scores for the first components (figure 5.12) show the same trend for both the LWN and the HWN window. Between the G1/S culture and the S phase culture, there is a slight yet statistically insignificant ($p = 0.10$ (LWN), $p = 0.29$ (HWN)) increase in nucleic acid and protein content relative to lipid content. There is no significant difference in the PCA scores between the S phase culture and the G2/M culture. However, between the G2/M culture and the early G1 phase culture, there is a significant decrease ($p = 0.0009$ (LWN), $p = 0.03$ (HWN)) in nucleic acid and protein content relative to lipid content. As was the case for the PCA scores for the first components from the asynchronous study (figure 5.4), the relative positions of the individual cell scores are consistent between the LWN and HWN windows. For example, cells #63 and #75 have, respectively, the highest and lowest scores from the 'G2/M +5 hrs' LWN window sample set, and the same two cells have, respectively, the highest and lowest scores from the corresponding HWN window sample set (figure 5.12).

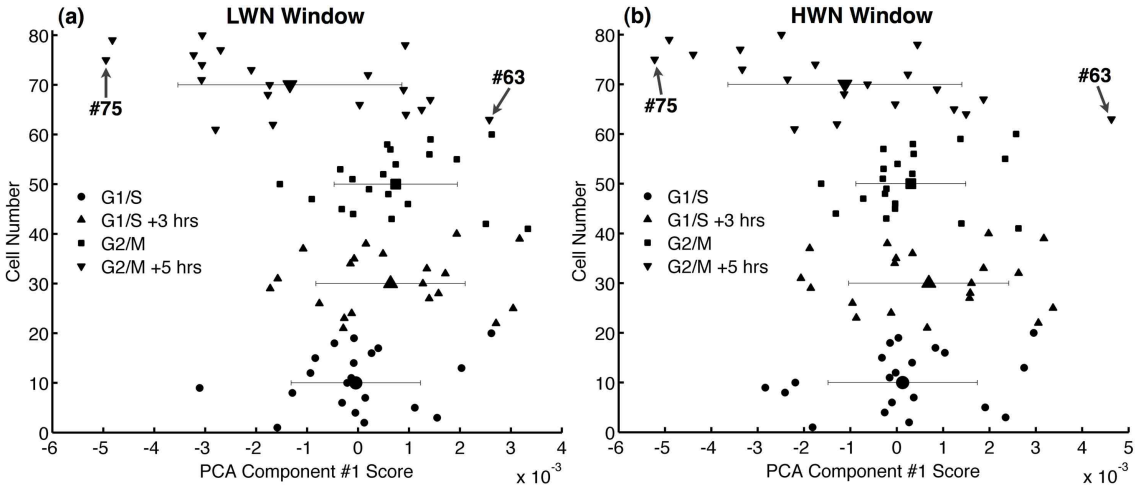


Figure 5.12: PCA scores for the first components from the synchronized cell cultures study, for the (a) LWN and (b) HWN window. The average score and standard deviation is shown for each sample for visualization purposes.

Second PCA component

The second component for the LWN window (figure 5.13a) explains 7.7% of the total variance, and the corresponding component for the HWN window (figure 5.13b) explains 2.1% of the total variance. Neither the LWN or HWN window components are similar to the second components obtained from the asynchronous cell cultures study (figure 5.6). For the LWN window all features are easily identifiable, except for the feature at 1402 cm^{-1}. The negative features include multiple contributions from the aromatic amino acids, with additional contributions from choline and O-P-O stretching in nucleic acids. The positive features are made up of contributions from nucleic acid bases and the DNA backbone, α-helix and β-sheet amide groups in proteins, and CH_2 twisting, C=C stretching, and both CH_2 and =CH deformation in lipids. For the HWN window, two broad negative features are observed, which possibly arise from the asymmetric stretching of CH_2 groups in lipids and the symmetric stretching of CH_3 groups in proteins and lipids.

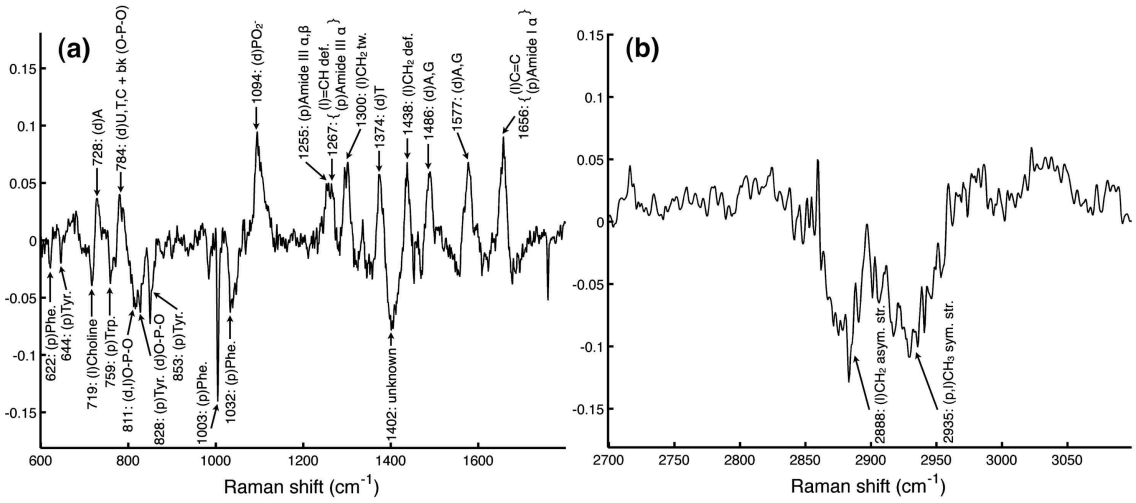


Figure 5.13: Second PCA components from the synchronized cell cultures study: (a) LWN window (7.7% of total variance), (b) HWN window (2.1% of total variance). The Raman shift and molecular origin of identifiable features are provided [59, 75, 92, 118–121]. Abbreviations - as in figure 5.1.

The PCA scores for the LWN window (figure 5.14a) show a significant increase ($p < 0.0001$) in the PCA scores for the G2/M culture. This increase is correlated with increased amounts of nucleic acid bases, DNA, conformational proteins, and CH_2 and C=C groups in lipids, and decreased amounts of aromatic amino acids, choline, and O-P-O groups in nucleic acids. The scores for the HWN window do not

have any relationship to the LWN window scores and do not appear to provide much meaningful biochemical information, except for that the highest scores are mostly observed in the early G1 phase culture.

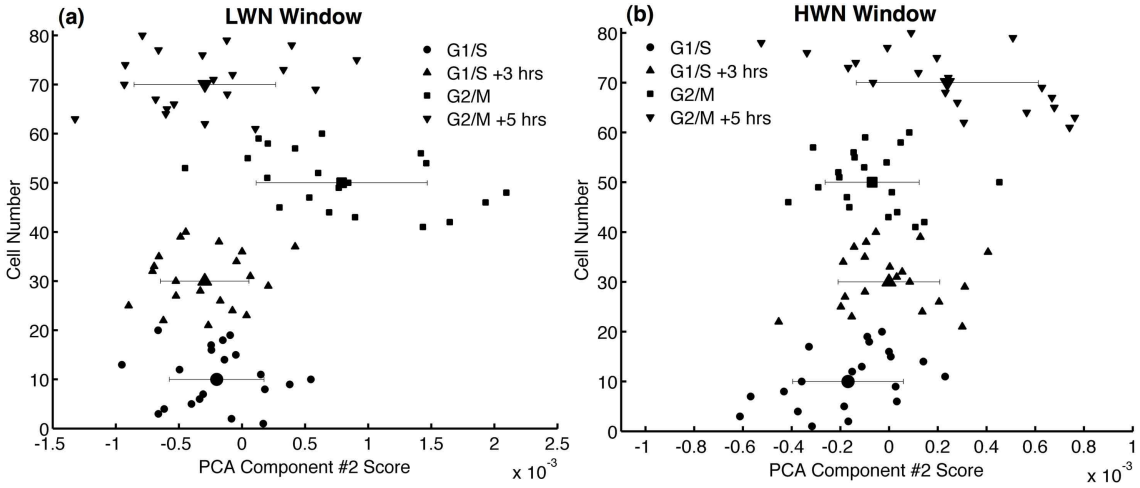


Figure 5.14: PCA scores for the second components from the synchronized cell cultures study, for the (a) LWN and (b) HWN window. The average score and standard deviation is shown for each sample for visualization purposes.

Other PCA components

The third PCA component for the LWN window (not shown) explains 5.1% of the total variance. Some features in this component are similar to those in the second PCA component from the asynchronous cell cultures study (figure 5.6a), including a strong negative contribution from choline at 719 cm^{-1} and a positive contribution from phenylalanine at 1003 cm^{-1}. However, the PCA scores for this component (also not shown) do not show any significant trend or discrimination between samples. The fourth and fifth components show features representative of slight X-axis calibration shifts, but since all the spectra in this study were collected in a single day the system calibration was very consistent for all samples; as such each component explains a small amount ($\sim$3%) of the total variance. Each of the remaining components for the LWN window explain less than 2% of the total variance and likely have little biological significance, and account for any residual variance arising from random sources of variability. The same can be said for all the remaining components for the HWN window, each of which explain less than 1% of the total variance.

5.4 Discussion

5.4.1 Study #1: Asynchronous cell cultures

The results of this 8-day study show that when Raman spectra are acquired from single DU145 cells taken from multiple cell cultures over multiple days, with different times between sub-culturing and Raman acquisition for each culture, there are primarily two independent sources of inherent variability observed in the Raman spectra. These two sources of variability are represented in this study by the first and second PCA components (figures 5.3 and 5.6).

First PCA component

For the entire 8-day data set in this study, the first PCA component explains 52.6% of the total variance for the LWN window data set. When searching for a biological origin for this component, an important consideration is that no matter which subset of the total 8-day data set is input into PCA, this same component is always observed as the primary source of variability, and typically explains 35-60% of the total variance. For example, if the data for only the first 4 days is input into PCA, the variance explained is 37.3%; however, if only the data for the last 4 days is used, the variance explained is 51.3%. No matter how many days worth of data are input into PCA, or which days are chosen, the primary features of the component do not change; namely the positive features arise from the same nucleic acid and protein molecules, and the negative features arise from the same lipid molecules, as those assigned to the first component in this study (figure 5.3a). These properties of the LWN component are also true for the HWN component (figure 5.3b), except that the percent variance explained is typically 75-90% of the total variance. It is also important to note that over the course of previous experiments (not shown), the Raman spectra from thousands of single DU145 cells have been collected and analyzed in this fashion. No matter which subset of the previously collected data is input into PCA, the first PCA components presented in this study (figure 5.3) are reproduced as the primary source of variability.

The most likely biological origin for the source of variability expressed by the first PCA components is the biochemical variability due to cell cycle (examined further below in the discussion of study #2). In this study, the PCA scores for the first PCA component, for both the LWN and the HWN window, show the trend of a steady

decrease in the average cellular nucleic acid and protein content relative to the average lipid content from 48 to 192 hours after sub-culturing (figure 5.4). Furthermore, there is a definite correlation between the steadily increasing fraction of cells in G1 phase as measured by flow cytometry (figure 5.2), and the steady decrease in the nucleic acid and protein content of individual cells relative to the lipid content, as measured by RS and calculated by PCA (figure 5.4). Interestingly, the flow cytometry results show that the fraction of cells in G1 phase begins to stop steadily increasing around 120-144 hours after sub-culturing, whereas the RS and PCA analysis shows that the relative nucleic acid and protein content continues to decrease steadily from 120 to 192 hours (figure 5.4). This discrepancy is likely in part due to a continual increase in the fraction of G1 cells that have entered into a non-proliferating quiescent 'G0' phase from 120 to 192 hours, which is not detectable by the methods used in this work. It is known that quiescent cells have a much lower RNA content than actively cycling G1 cells, as well as a decreased amount of certain proteins required for cell cycle progression [17, 18]. The discrepancy is also likely in part due to the flow cytometry observation that the fraction of cells in S phase continues to decrease from 120 to 192 hours, and reaches a minimum of $\sim$8% at 192 hours after sub-culturing. However, an increased fraction of quiescent cells and a decreased fraction of S phase cells are both indicators of a less proliferative cell culture, which is an expected trend as cells are left for longer periods of time after sub-culturing.

It is important to note that the results of this study are in agreement with two previous Raman studies [83, 123], both which investigated spectral changes due to cell cycle by comparing averaged bulk Raman spectra from "exponential phase" cell cultures (G1 fraction $\leq$ 50%) to spectra from "plateau phase" cell cultures (G1 fraction $\geq$ 80%). One of these studies [83] found that protein/lipid, RNA/lipid and DNA/lipid ratios are all statistically higher for exponentially growing cells, as determined by fitting biochemical component spectra to the measured LWN and HWN window averaged Raman spectra. This same study also identified the spectral regions, and the corresponding molecules assigned to those regions, which yielded significant averaged spectral differences between samples. With similar methods, the other study [123] demonstrated that increased fractions of both protein and nucleic acid content in exponentially proliferating cells is correlated with decreased fractions of lipid and glycogen content, as compared to plateau phase cells. The results presented here on RS of single cells corroborate and extend these previous Raman results for bulk samples. The results of the asynchronous cell cultures study identify the individual

molecular sub-groups that are most responsible for the observed changes in Raman spectra, such as the strong contribution from CH_2 deformation in lipids in the first PCA component for the LWN window (figure 5.3a). The PCA results presented here also demonstrate that changes in the relative lipid content in a cell are mathematically anti-correlated with changes in both the protein and nucleic acid content in a cell; this result is consistent both with previous Raman results [123] and with the prior knowledge that the RNA-to-protein ratio is relatively constant within a cell throughout the cell cycle [17]. The results extend previous Raman studies by showing that the changes in biochemical composition due to cell cycle can be directly observed in single-cell spectra (*i.e.,* figure 5.5), and that the changes can be readily observed as a continuous process as a cell culture moves from an exponentially growing culture (24 to 96 hours after sub-culturing) to a confluent non-exponential culture (120 to 192 hours after sub-culturing). Finally, as discussed below, this study demonstrates that there is another significant source of variability (arising from differences in cell culture confluency) that is detectable when performing RS on single cultured DU145 cells, in addition to the variability in the nucleic acid and protein content relative to lipid content.

Second PCA component

For the entire 8-day data set in this study, the second PCA component explains 10.1% of the total variance for the LWN window data set. However, unlike the first PCA component, the amount of variance explained by this component is highly dependent on which subsets of the total data set are input into PCA. For example, the variance explained is maximized at 16.7% when only the data for the first 5 days is input into PCA. However, when the data for the first 2 days is excluded the variance explained drops from 10.1% to 4.8%, and when the first 3 days are excluded the variance explained drops further to 2.4%. When the first 4 days, or more, are excluded, the variance explained becomes less than 2% and the component is no longer recognizable. These properties of the LWN component are also true for the HWN component (figure 5.6b); in the HWN case the variance explained is maximized at 3.3% when only the data for the first 5 days is input into PCA, yet the component is not observed when the first 4 or more days are excluded, as was the case for the LWN component. The dependency of the second PCA component on the choice of sample subset is consistent with the corresponding PCA scores (figure 5.7), which steadily increase up to 5 days after sub-culturing and remain fairly constant from 5 to 8 days after sub-culturing.

A biological explanation for the spectral variability described by the second PCA component is unclear, especially for the HWN window where the molecular origins of the features are unknown (figure 5.6b). However, there is a strong correlation between the trend of the PCA scores (figure 5.7) and the measured confluency of the cell cultures (figure 5.2), which is in turn related to the amount of time the culture was left to incubate after sub-culturing. In this study, the cell cultures steadily increase in confluency until ∼5 days after sub-culturing, after which there is very little room left to grow and the confluency remains relatively constant at ∼90%. The confluency trend matches the trend of the PCA scores, which steadily increase up to 5 days after sub-culturing and remain fairly constant afterwards. Furthermore, as discussed above, if only the data from days 5 to 8 is input into PCA (*i.e.*, only the data collected once the culture had reached ∼90% confluency), then the second PCA component presented here is not observed at all. This implies that the component is directly caused by biochemical changes related to a sub-confluent culture growing during the first 3 to 4 days after sub-culturing. The primary biomolecules responsible for this source of variability in the Raman spectra can be identified in the second PCA component for the LWN window (figure 5.6a). This study is the first spectroscopic measurement of inherent biochemical variability in a cell culture that is correlated with the changing confluency of a cell culture during the first 3-4 days after sub-culturing, which is shown here to be independent of the existing variability arising from differences in cell cycle progression. This source of variability has been observed in several previous experiments with DU145 cells (not shown), in which cell cultures were harvested for Raman analysis 1-2 days after sub-culturing. It should be noted that although this source of variability is found to be significant for DU145 cells, it is not a characteristic of all *in vitro* cultured cell lines, as observed in the studies presented in chapter 7.

5.4.2 Study #2: Synchronized cell cultures

In theory, the effects of the cell cycle regulatory drugs thymidine and nocodazole are reversible, such that when the drug is removed and replaced by fresh media the cells in the culture progress synchronously through the rest of their mitotic cycle. In practice, whole culture synchronization is unfeasible and cell cultures become desynchronized within ∼24 hours [175, 176]. Furthermore, some fraction of the culture will not be immediately released (or released at all) from the drug induced arrest. However,

large fractions of cells (*i.e.*, ∼75%) can indeed be synchronized for short periods of time (typically less than 24 hours), and the cell cycle distribution can therefore be significantly altered from that of an untreated asynchronous population. Despite the known shortcomings of using drugs for cell cycle control, drug treatment is still the easiest and simplest technique for significantly altering the cell cycle distribution and achieving a high yield of mostly synchronized cells. It is important to note that the level of synchronization reported in this study (figure 5.10) is comparable to the level of synchronization reported by both the recent study using RS for cell cycle discrimination, where synchronization was performed by serum starvation and cell cycle regulatory drugs [84], and another previous study that investigated the infrared spectroscopic differences between cells in different stages of the cell cycle, where synchronization was performed by centrifugal elutriation [177].

A main goal of this second study is to directly examine whether the variability expressed by the first PCA components (addressed above in the discussion for study #1) is indeed due to biochemical differences between cells at different points in the cell cycle. The first PCA components for this study (figure 5.11) have primarily the same features as the first PCA components for the asynchronous cell cultures study (figure 5.3), and both the LWN and HWN window components explain approximately the same amount of the total variance as the corresponding components in the asynchronous cell cultures study. However, in this study all four cultures were harvested on the same day (after each were synchronized) and the confluency of each culture was <60%; as such, all four cultures should have very few quiescent cells, and one would not expect to measure any variability due to the differences in confluency between the cultures.

The variability in the nucleic acid and protein content relative to the lipid content in single cells, between the four synchronized cell cultures, is expressed by the PCA scores for the first components (figure 5.12). The slight shift to a higher average relative nucleic acid and protein content between the 'G1/S' and the 'G1/S +3 hrs' cultures correlates with the flow cytometry measured shift from ∼83% of the first culture existing at the G1/S boundary to ∼64% of the second culture progressing through S phase (figure 5.10). This shift is consistent with expected changes in the biochemical content for S phase cells, which contain increased levels of RNA and protein as compared to G1 cells [17], and an increased amount of DNA due to the active DNA replication that occurs during S phase. There is no observable shift in the scores between the 'G1/S +3 hrs' and the 'G2/M' cultures (figure 5.12), even though

the 'G2/M' culture has over 74% of its cells at the G2/M boundary compared to only ∼16% of the cells in G2 phase for the 'G1/S +3 hrs' culture. This lack of separation in the scores for the first components may seem at odds with a known increase in the overall RNA and protein content of G2/M cells as compared to late S phase cells [17]; however, the PCA scores for the first components only represent changes in nucleic acid and protein content relative to the total lipid content, which is also increasing throughout G2 phase in preparation for mitotic division. The most significant change in the scores for the first PCA components occurs as a decrease in the average nucleic acid and protein content between the 'G2/M' and the 'G2/M +5 hrs' cultures, which undergo a transition from a culture with >74% of its cells in a G2/M phase to a culture with ∼75% of its cells existing within the first 5 hours of G1 phase. The observed change in the relative nucleic acid and protein content is consistent with previous biochemical experiments, which have shown that the lowest levels of RNA and protein are found within the first few hours of G1 phase, immediately following cell division [17, 18]. These observations are also consistent with the results of the recent RS study for cell cycle discrimination [84], where the successful discrimination between S or G2/M cells and G0/G1 cells was due to increased nucleic acid and protein content relative to lipid content in both S and G2/M cells, as measured in the LWN spectral window. This previous study also reported poor discrimination between S and G2/M cells based on nucleic acid and protein content relative to lipid content [84], which is also observed here for both spectral windows (figure 5.12). In the study presented here, it is interesting to note that the PCA scores for the LWN window (figure 5.12a) for the 'G2/M +5 hrs' culture are well split into two subgroups: 60% of the cells have PCA scores < -1 (low relative amount of nucleic acid and protein) and 40% of the cells have PCA scores > 0 (high relative amount of nucleic acid and protein). This split is matched by the cell cycle distribution for this culture (figure 5.10), which is distinctly separated into two groups: ∼75% of the cells in early G1 phase and ∼21% of the cells in G2 phase, with only ∼4% of the cells in S phase. The relative positions of the scores for this culture are similar for the HWN window, but the separation between the two subgroups is less distinct (figure 5.12b). In summary, these results confirm that the most significant source of Raman spectral variability between cells in a culture, which is expressed in this work by the first PCA component from each spectral window, can be confidently attributed to biochemical changes arising from the progression of individual cells through their mitotic cycle.

The features in the second PCA components for this study (figure 5.13) are dif-

ferent than the features in the second PCA components for the asynchronous cell cultures study (figure 5.6). The scores for the LWN and HWN windows do not show similar trends, which suggests that the components arise from different sources of biochemical variability between cultures. The scores for the HWN component (figure 5.14b) do not show any clear trend or separation between cultures, and the features in the HWN component (figure 5.13b) are not clearly attributed to a unique set or class of biomolecules. As such, it is difficult to assign a biological meaning to the HWN window results. However, the scores for the LWN component (figure 5.14a) distinctly separate the 'G2/M' cells from the other cultures with an increase in the average PCA score. According to the corresponding PCA component (figure 5.13a), the increase in scores for the G2/M cells corresponds primarily with a decrease in aromatic amino acids, choline and O-P-O groups in RNA and/or lipids, correlating with an increase in nucleic acid bases, DNA, α-helix and β-sheet amide groups, and CH_2, =CH and C=C lipid groups. The biological reason for these changes is unclear, but may be related to changes in the cell biochemistry in preparation for mitotic division. Alternatively, the variability could arise as a temporary cellular response to the nocodazole treatment for synchronization of the G2/M culture. If these measured changes are indeed due to a natural source of variability inherent to G2/M cells, it is likely that this component is not observed in the asynchronous cell culture study due to the low fraction of G2 cells in each of the asynchronous cultures, and the presence of other larger sources of variability such as the changing confluency and shifts in the X-axis calibration. Further work using different synchronization techniques would be required to determine if this source of variability is indeed an inherent characteristic of G2/M cells.

5.4.3 Spectral variability and PCA

Many of the results in this work depend on the accurate assignment of a molecular origin to features in the PCA components. However, achieving confidence in the validity of such molecular assignments for the features in the PCA components is only possible if all external sources of spectral variability, which are not inherent to the biomolecular composition of the cells, are removed prior to PCA implementation. Sources of variability that may arise include variability in the intensity and shape of the baseline (originating from a sample substrate, water, or from the cellular material itself), or variability induced into a data set by an unsuitable spectral

normalization technique. If an external source of variability happens to contribute preferentially to a certain sample (or samples) in the data set, the PCA algorithm will faithfully correlate the variability from the external source with any variability that is inherent to the sample in question, which is possibly the variability of interest in the experiment. An example of this issue, which arose during the course of this work, is the variability in intensity of spectral contributions from the quartz substrate. For the 'G2/M +5 hrs' culture from the synchronized cell culture study, the pellet of cells used for Raman acquisition was only a few cell layers thick, as opposed to tens of cell layers thick for the other cultures in the study. As such, there was a slightly greater quartz contribution observed in all the spectra collected from the 'G2/M +5 hrs' cells. If a very conformal baseline was not applied for baseline estimation in the lower half of the LWN window ($\sim$600-1200 cm^{-1}), it was found that the variability due to quartz became significant to the degree that the quartz variability became correlated with other sources of variability inherent to the 'G2/M +5 hrs' culture. As the conformity of the baseline was reduced, the resultant increase in quartz variability became observable in the first PCA component as recognizable quartz features, and many biological features from the first PCA component began to appear in the component that was originally dominated by quartz features alone. These considerations require extreme care when developing and implementing automated spectral processing methods, such as spectral smoothing or baseline estimation methods, when large multi-sample data sets are prepared for PCA implementation. However, the PCA components themselves can aid in the identification of external sources of variability during the development and implementation of spectral processing techniques (as shown in chapter 4), as long as the spectral features of the external sources are known. It should be noted that performing Raman analysis with the HWN window is significantly simpler, since baseline removal is easier due to the absence of substrate contributions and fluorescence in this spectral region. However, the simplicity advantage comes at the cost of a significant decrease in the amount of biochemical information available, as compared with the LWN window. Furthermore, the spectral contributions from water in the HWN window may become significant if the methods presented here are applied to the Raman analysis of cells in an aqueous environment.

5.4.4 LWN vs. HWN spectral windows

Both the LWN and HWN windows are analyzed independently throughout this work to determine if information can be obtained equivalently from either window. It is found that biochemical variability due to cell cycle is clearly observable in either window, and the spectral differences are directly observable in the original data for both windows (figures 5.5a and 5.5b). However, the LWN window provides information from many more biomolecules, including multiple contributions from nucleic acids and amino acids that are not observed as sources of cell cycle variability in the HWN window. The variability due to changes in cell culture confluency after sub-culturing is more apparent in the LWN window, due to the strong contributions from the features identified in the second PCA component from the asynchronous cell cultures study (figure 5.3a). Although the trends of the PCA scores for the second components are similar for both the LWN and HWN windows (figure 5.4), the molecular origins of the corresponding features in the HWN window component are uncertain (figure 5.3b). Furthermore, the spectral differences arising from this source of variability are directly observable in the original data only for the LWN window (figure 5.8). Therefore, in this case the LWN window provides spectroscopic information that is not available in the HWN window. It is also shown that the LWN window is sensitive to biochemical changes unique to the G2/M sample from the synchronized cell cultures study, whereas in the HWN window no spectroscopic differences are observed for the same sample.

5.4.5 Spectral variability and cell size

All of the single-cell RS measurements presented here are acquired with a fixed sampling volume ($\sim$2 x 5 x 10 μm, in x-y-z), which is aligned with the centre of the selected cell (recall figure 3.5). Therefore, there is the possibility of observing spectral differences that correlate simply with size differences in the cell population. For example, previous authors [80] have noted that a smaller cell has a higher surface area to volume ratio than a larger cell, and may therefore yield more biochemical signals from cell membrane lipids and proteins relative to cytoplasmic and nuclear biomolecules. However, these authors used direct measurements of the size of each selected cell, obtained in suspension during optical tweezers RS acquisition, to show that cell size had no correlation with the ability of RS and PCA to biochemically discriminate between two cell lines of different average size [80]. In this work, moni-

toring absolute cell size via direct optical measurements of the cells selected would be inaccurate due to the lack of 3D cell geometry information in the pellet. However, the relative cell size distribution for a given culture is well described by the flow cytometry measurement of forward scatter intensity, which is acquired from cells in suspension simultaneously with PI fluorescent intensity acquired for cell cycle analysis. To verify that the dominant sources of inherent spectral variability observed in this work (as described by the first and second PCA components in study #1) are not simply due to changing cell size, the forward scatter intensity distributions of all 8 cell cultures used in study #1 were inspected, and the changes between cultures were compared to the trends of the PCA score distributions for each PCA component. The results of this analysis are summarized below.

From 24 to 72 hours after sub-culturing, there is no detectable change in the cell size distribution between cultures. From 72 to 96 hours, there is a detectable shift in the measured distribution towards lower forward scatter intensities, which is indicative of a higher proportion of smaller cells in the culture.* From 96 to 192 hours after sub-culturing, there is no further detectable change in the cell size distribution between cultures. However, the PCA results presented here indicate that the largest source of spectral variability observed in this study (first PCA component) displays a steady trend of continuing spectral differences occurring from 48 to 192 hours after sub-culturing (figure 5.4). Furthermore, the second largest source of spectral variability (second PCA component) displays a trend of continuing spectral differences occurring from 24 to 120 hours after sub-culturing (figure 5.7). Neither of these sources of variability correlate with the observed changes in the relative cell size distribution. Therefore, any spectral variability arising from differences in cell size must be explained by one of the many lower variance PCA components, each of which explain less than 3% of the total variance for the LWN window, and less than 1% of the total variance for the HWN window. Differences in cell size may introduce significant spectral variability when comparing cell lines with large differences in average size, but within a single cell line this analysis shows that cell size variations are not a significant source of spectral variability.

*This shift is consistent with the cell cycle analysis (figure 5.2) where between 72 and 96 hours there is a sharp increase in the fraction of G1 phase cells, which are typically smaller than S phase and G2 phase cells.

5.5 Conclusion

The studies in this chapter characterize the inherent sources of spectral variability between single cells of a human prostate tumour cell line (DU145) cultured *in vitro*. PCA is used to identify spectral differences that correlate with cell cycle progression and the changing confluency of a cell culture during the first 3-4 days after sub-culturing. Spectral variability arising from cell cycle progression is (i) expressed as varying intensities of protein and nucleic acid features relative to lipid features, (ii) well correlated with known biochemical changes in cells as they progress through the cell cycle, and (iii) shown to be the most significant source of inherent spectral variability between cells. The results presented in this chapter provide a foundation for interpreting spectral variability in subsequent studies, and demonstrate the levels of sensitivity and specificity achieved by the RS acquisition and data processing techniques developed in this work.

Chapter 6

Results & Discussion III: Raman spectroscopy of single human tumour cells exposed to ionizing radiation

6.1 Introduction

This chapter presents work published by IOP Publishing (Bristol, UK) in the journal *Physics in Medicine and Biology*, entitled "Raman spectroscopy of single human tumour cells exposed to ionizing radiation *in vitro*" [178] (doi:10.1088/0031-9155/56/1/002), and is reproduced here with permission from IOP Publishing.

This studies presented in this chapter investigate the capability of Raman spectroscopy (RS) to study the effects of ionizing radiation on single human tumour cells. As discussed in section 1.3.3, previous studies have successfully used RS to detect spectral changes in irradiated aqueous DNA [85, 86], biological membranes [61, 88–91], and skin and muscle tissues [92, 93]. In addition, a recent study applied RS to discriminate between responding and non-responding cervical cancers after irradiation [94], although the authors were unable to draw any conclusions regarding the differences in the biochemical composition of the tissues. To date, Raman techniques have yet to be applied to investigate the radiation response of single living cells, which are a vital experimental source from which many advances in radiobiology arise.

In this chapter, human prostate tumour cells (cell line DU145) are cultured *in vitro* and irradiated to doses between 15 and 50 Gy with single fractions of 6 MV photons.

Irradiated and unirradiated cell cultures are re-incubated for varying amounts of time post-irradiation, up to five days. For each cell, Raman spectra are obtained from both the LWN region and the HWN region in order to determine the utility of each spectral window in detecting radiation induced spectral changes. PCA is used to show that Raman spectra collected from irradiated cells display a measurable radiobiological effect that is correlated with both dose and post-irradiation incubation time. The effect is expressed as varying concentrations of specific molecular groups, assigned to lipids, nucleic acids, amino acids and conformational protein structures, within irradiated cells as compared to unirradiated cells. PCA is shown to be useful in discriminating between radiation induced changes in cell spectra and inherent spectral differences within a cell culture arising from cell cycle progression and other factors, which are examined in detail in chapter 5. The results of two separate irradiation experiments are presented in this chapter (sections 6.3.2 and 6.3.3), the latter which investigates the effect of changing the time of irradiation after the initial preparation of the cell cultures.

6.2 Methods

The general procedure followed for the irradiation experiments presented below was to prepare a number of identical cell cultures, and irradiate each cell culture at the same time while leaving some cultures unirradiated to serve as controls. The remaining live cells were harvested from selected cell cultures for RS at varying time points (0 to 5 days) post-irradiation, as described in section 3.1.4. For the first experiment presented below in section 6.3.2, cells were allowed to incubate for $\sim$3.5 days prior to irradiation in order to avoid measurement of any additional inherent spectral variability correlated with the incubation time of DU145 cell cultures during the first 3-4 days after sub-culturing (as shown in chapter 5). Spectra were collected from 20 individual cells from 12 cell cultures, according to the protocol depicted in figure 3.2, resulting in 240 spectra collected for each spectral window. The differences in protocol for the experiment presented in section 6.3.3, which investigates the effect of changing the time of irradiation after sub-culturing, are provided therein.

Prior to RS analysis, the cell cycle distribution and viability of each culture was determined with flow cytometry, as described in previous chapters. For the experiment presented in section 6.3.2, the fraction of live cells in the harvested samples was over 95% for all cultures, with the exception of the irradiated 120 hours sample which

had a live cell fraction of ∼90%.

RS acquisition and data processing was performed as described previously. No spectral variability arising from the quartz substrate was observed in the experiments in this chapter; as such, no filtering was performed with PCA.

6.3 Results

6.3.1 Single DU145 cell spectrum

The LWN and HWN spectral window Raman spectra for a single unirradiated DU145 cell from this work are shown in figure 6.1. The spectral features and molecular assignments observed in the single DU145 cell spectra presented in the previous chapter (figure 5.1 and table 5.1) are reproduced in the studies presented here, and the spectra are provided herein (figure 6.1) for ease of reference.

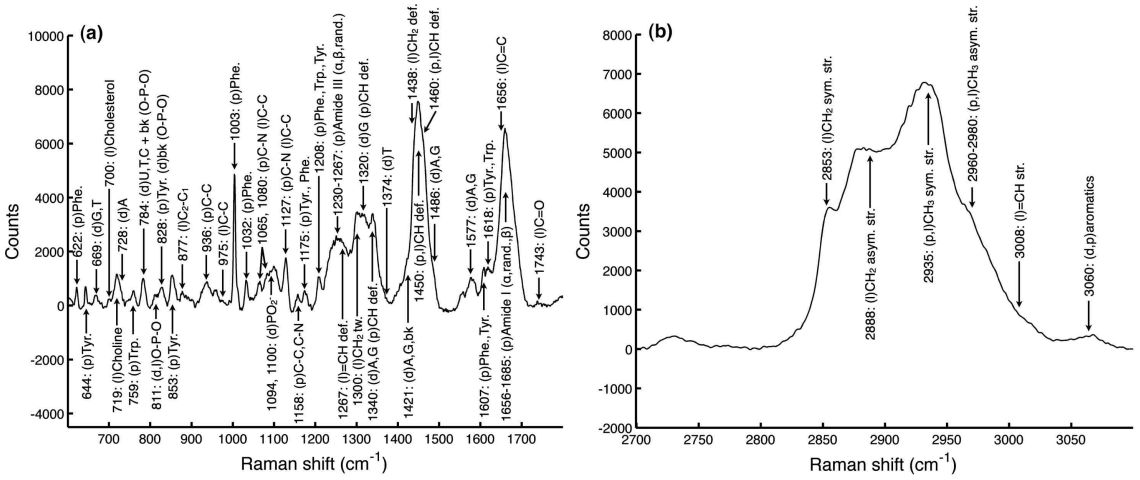


Figure 6.1: Raman spectra of a single unirradiated DU145 cell, for the (a) LWN and (b) HWN spectral windows; the Raman shift and molecular origin of identifiable features are provided [59, 75, 92, 118–121]. Abbreviations - p: protein, l: lipid, d: DNA/RNA, A: adenine, T: thymine, G: guanine, C: cytosine, U: uracil, Phe: phenylalanine, Tyr: tyrosine, Trp: tryptophan, bk: backbone, def: deformation, tw: twist, sym: symmetric, asym: asymmetric, str: stretch.

6.3.2 Irradiated vs. unirradiated cells

LWN window: First PCA component

As described in section 6.2, a total of 240 LWN window cell spectra were collected over the course of this experiment. The first PCA component (figure 6.2a) describes the most significant source of spectral variability, which in this case is 48.8% of the total variance in the data set. As before (chapter 5), molecular origins can be assigned to many of the positive and negative features in the component (figure 6.2a). The positive features in the component are exclusively nucleic acid and protein in origin, with nucleic acid features arising from DNA and RNA bases and the DNA backbone, and protein features arising from aromatic amino acids (phenylalanine, tryptophan and tyrosine) and β-sheet amide groups. Conversely, the negative features in the component are dominated by lipid contributions from cholesterol, choline, CH_2 twisting, CH_2 and CH deformation, and C-C, C=C and C=O stretching, with some overlap of protein features at 1065, 1080, 1267, 1460, and 1656 cm^{-1} arising from C-N stretching, CH deformation and α-helix amide groups. It cannot be determined whether the variability described by these overlapping negative features is uniquely lipid or protein in origin, but since the rest of the negative features in the component are uniquely lipid, it is likely that lipids are contributing in part to the negative nature of these features as well. To summarize, the positive features in the first PCA component arise from nucleic acids and proteins, whereas the negative features arise primarily from lipids.

Recall that any spectrum with a higher (*i.e.*, more positive) PCA score for a given PCA component contains a proportionately higher amount of the positive features, and lower amount of the negative features, from that component. For the first PCA component, positive scores are correlated with increased nucleic acid and protein content, and negative scores are correlated with increased lipid content. The PCA scores for the 240 spectra in this experiment (figure 6.2b) are grouped by time of RS acquisition and by irradiated dose for each sample. For the unirradiated cultures, the average nucleic acid and protein content relative to the average lipid content stays relatively constant until decreasing from 48 to 72 to 120 hours after time of irradiation. For the irradiated cultures, the average nucleic acid and protein content relative to the average lipid content is consistent with the unirradiated cultures at 0 and 24 hours post-irradiation, decreases by 48 hours post-irradiation, and remains fairly constant from 48 to 120 hours post-irradiation. Overall, the PCA scores for the first PCA

component show no clear separation between the irradiated and unirradiated cell cultures as a function of dose or incubation time post-irradiation.

In chapter 5 (and in the corresponding publication [168]) it is shown that the variability expressed by this first PCA component during RS analysis of single DU145 cells arises from inherent cell cycle variability between cells, both within a given culture and between different cultures (discussed further in section 6.4.1). In the current work, the observed changes in the average nucleic acid and protein content relative to the average lipid content of unirradiated cells as a result of different incubation times after sub-culturing (figure 6.2b) are consistent with previous observations, and are also consistent with known changes in the biochemical composition of cells in cultures allowed to incubate for longer times after sub-culturing [17, 18, 83]. In all studies performed to date with irradiated and unirradiated DU145 cell cultures, cell cycle variability has consistently remained the largest source of spectral variability observed within a given data set (typically 40-60% of the total variance), and is often easily identifiable in the original spectra. To demonstrate this last point, the Raman spectra and Raman difference spectrum for two unirradiated cells (cells #41 and #214) with a large separation in PCA score (figure 6.2b) are shown in figure 6.2c, along with the first PCA component for comparison. Nearly all of the features previously identified in the component (figure 6.2a), which describes the spectral variability due to cell cycle in the entire data set, are directly observable in the difference spectrum of two cells with a large separation in PCA score for the first PCA component.

LWN window: Second PCA component

The second PCA component (figure 6.3a) describes the second most significant source of spectral variability in the data set (14.4% of total variance), which is independent of the existing variability due to cell cycle described by the first PCA component (figure 6.2). As before, by comparing the features in the component with known molecular Raman shifts (figure 6.1a), molecular origins can be assigned to many of the positive and negative features in the component. The positive features include contributions from aromatic amino acids, C-C and C-N stretching and CH deformation in proteins, CH_2 twisting and CH and CH_2 deformation in lipids, and a combined contribution from the nucleic acid bases A and G and CH deformation in proteins. The negative features include contributions from choline, O-P-O stretching in lipids and RNA, C-C stretching in lipids, random coil amide groups, the nucleic acid bases and the DNA

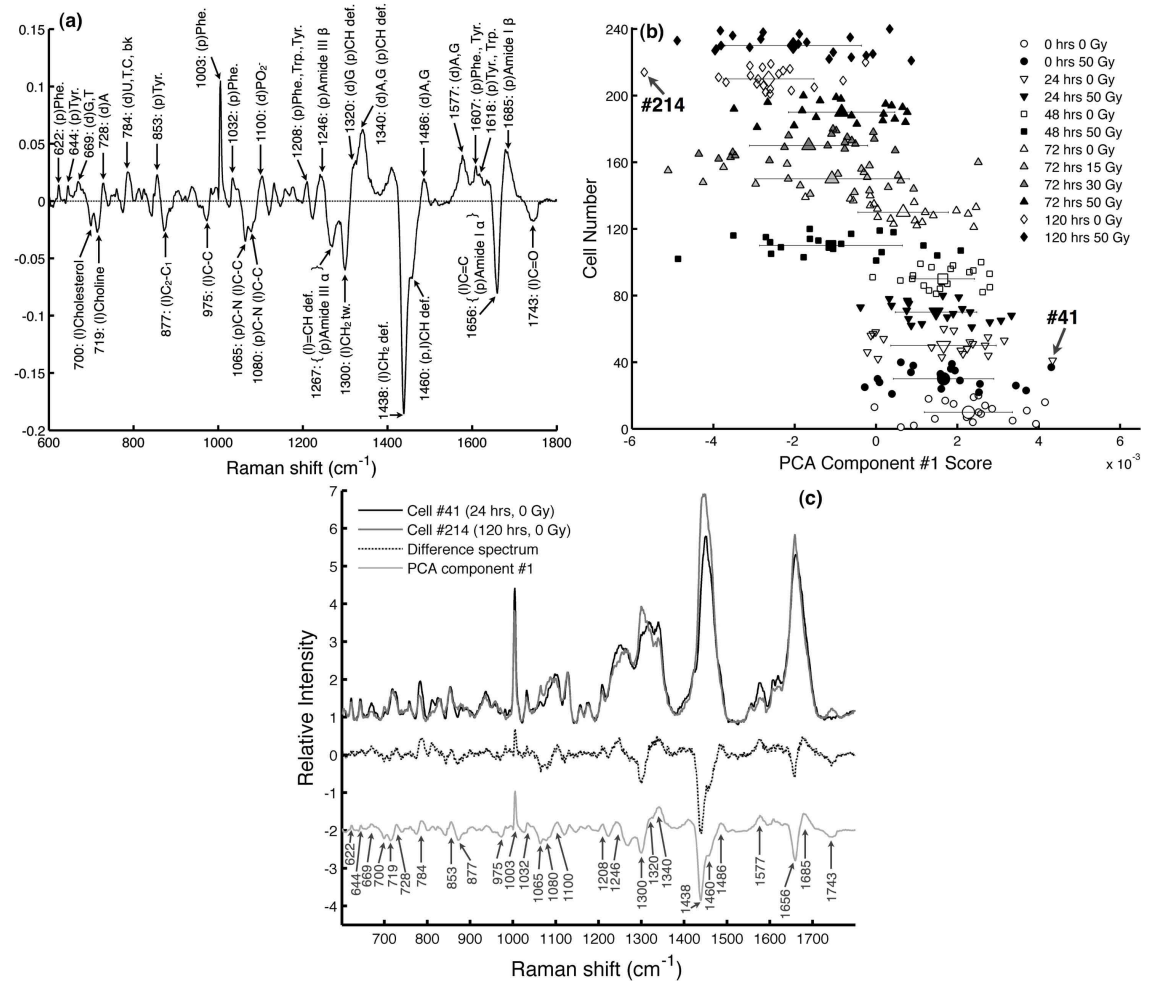


Figure 6.2: First PCA component results for the LWN window. (a) First PCA component (48.8% of total variance), with Raman shifts and molecular origins (figure 6.1a) of identifiable features. Abbreviations - as in figure 6.1. (b) PCA scores for the first PCA component. Different markers categorize all 240 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data. (c) Raman spectra and Raman difference spectrum for two unirradiated cells (#41 and #214) with a large difference in PCA score. Also shown is the first PCA component, which has been offset and rescaled for comparison with the unscaled difference spectrum. Wavenumbers are provided for any known features in the component which are observable in the difference spectrum.

backbone, and a combined contribution from lipid =CH deformation and α-helix amide groups. The sharp negative feature at 1660 cm^{-1} arises from amide groups as well, and is likely a combined contribution from α-helix (1656 cm^{-1}) and random coil (1669 cm^{-1}) amide groups. Since the negative features from the nucleic acid bases A and G at 1486 and 1577 cm^{-1} have no overlapping contributions from other molecules, it is likely that the positive nature of the combined feature at 1340 cm^{-1} is from CH deformation in proteins (which also contributes positively at 1460 cm^{-1}) rather than from the nucleic acid bases A and G. The positive feature at 1421 cm^{-1} is tentatively assigned to CH_2 groups in the backbone of nucleic acids, but the accuracy of this assignment is uncertain due to contradicting literature sources [75, 92, 119, 120].

The PCA scores (figure 6.3b) for the second PCA component display no separation between the 50 Gy and 0 Gy cell cultures immediately after irradiation ($p = 0.15$). However, the average PCA scores show a clear separation after 24 hours ($p < .0001$), which increases steadily from 24 to 120 hours post-irradiation. Furthermore, at 72 hours post-irradiation there is a steady decrease in the PCA scores as the dose increases from 0 to 15 to 30 to 50 Gy ($p < .0001$ for each increase in dose). These trends in the PCA scores show that the biochemical changes described by the second PCA component (figure 6.3a) are correlated with both the incubation time post-irradiation and the irradiated dose after a fixed incubation period.

In this study the radiation induced spectral variability explains much less of the total percent variance than the inherent spectral variability due to cell cycle (14.4% for radiation vs. 48.8% for cell cycle). As such, it is difficult to directly observe radiation induced spectral changes in the Raman difference spectrum from two cells. However, if the chosen cells have a large difference in PCA score for the second PCA component and a small difference in PCA score for the first PCA component (*i.e.*, large spectral differences caused by radiation and small spectral differences due to cell cycle), the radiation induced spectral changes are clearly visible in the Raman difference spectrum (figure 6.3c).

LWN window: Other PCA components

The third PCA component for the LWN window (figure 6.4a) explains 4.7% of the total variance. The features in the component are dominated by derivative-like features centred at the positions of the sharpest and steepest Raman features in the spectral window, most notably at the phenylalanine ring breathing peak at 1003 cm^{-1}. These features in a PCA component indicate variability arising from shifts in the calibra-

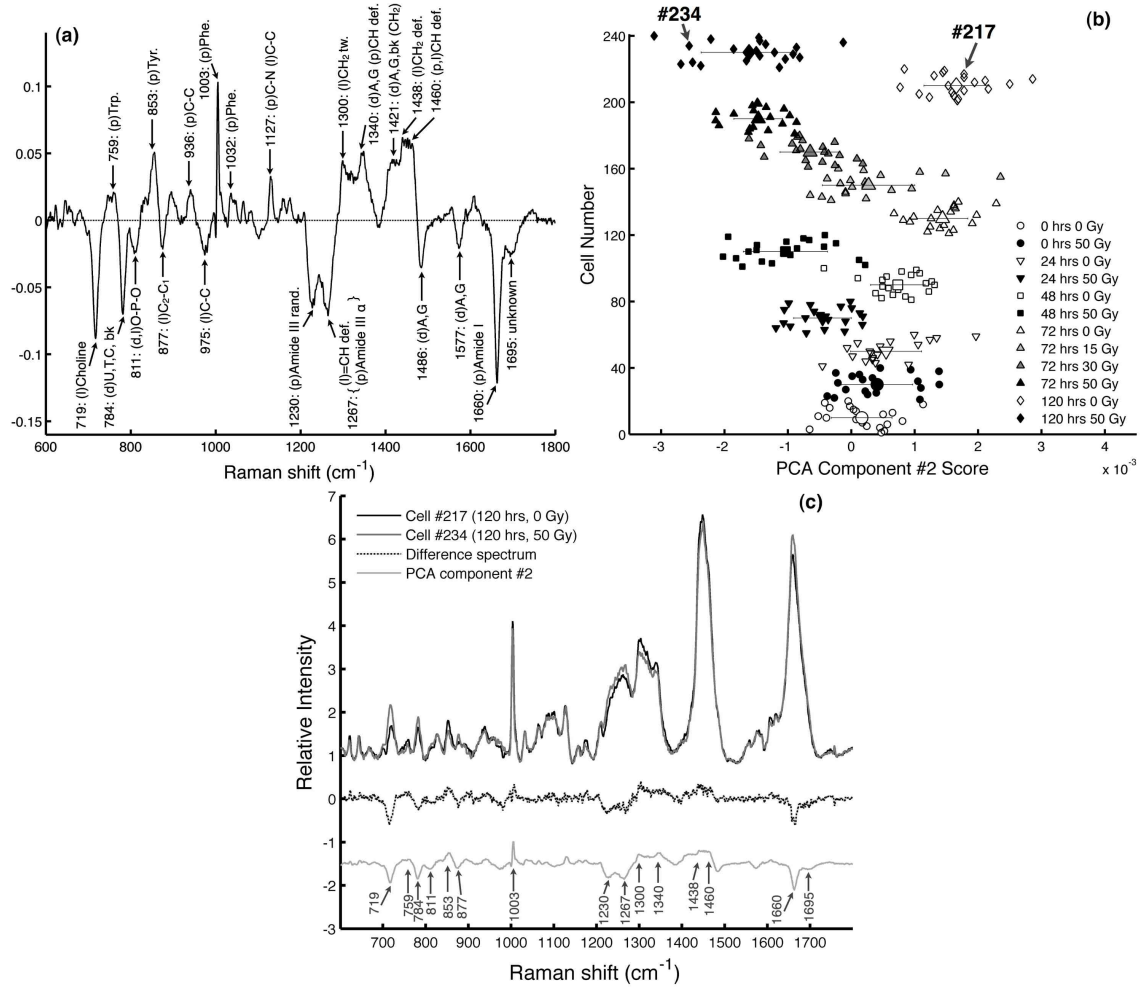


Figure 6.3: Second PCA component results for the LWN window. (a) Second PCA component (14.4% of total variance), with Raman shifts and molecular origins (figure 6.1a) of identifiable features. Abbreviations - as in figure 6.1. (b) PCA scores for the second PCA component. Different markers categorize all 240 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data. (c) Raman spectra and Raman difference spectrum for two cells (#217 and #234) with a large difference in PCA score for the second PCA component and a small difference in PCA score for the first PCA component. Also shown is the second PCA component, which has been offset and rescaled for comparison with the unscaled difference spectrum. Wavenumbers are provided for any known features in the component which are observable in the difference spectrum.

tion of the Raman system. The trend displayed by the score plots for this component (figure 6.4b) correlates well with measured shifts in the Raman calibration over the 5 day sample collection period, which were monitored by measuring the peak position of the 520 cm^{-1} feature of the instrument-based silicon sample before and after RS collection each day. Before RS collection, it was first verified that the initial calibration of the system was within 0.5 cm^{-1} of the calibration performed on the first day of collection, which translates to a maximum shift in the system calibration of less than one pixel (1 pixel $\approx$ 0.9 cm^{-1}, at 1003 cm^{-1}) for all 240 spectra collected.

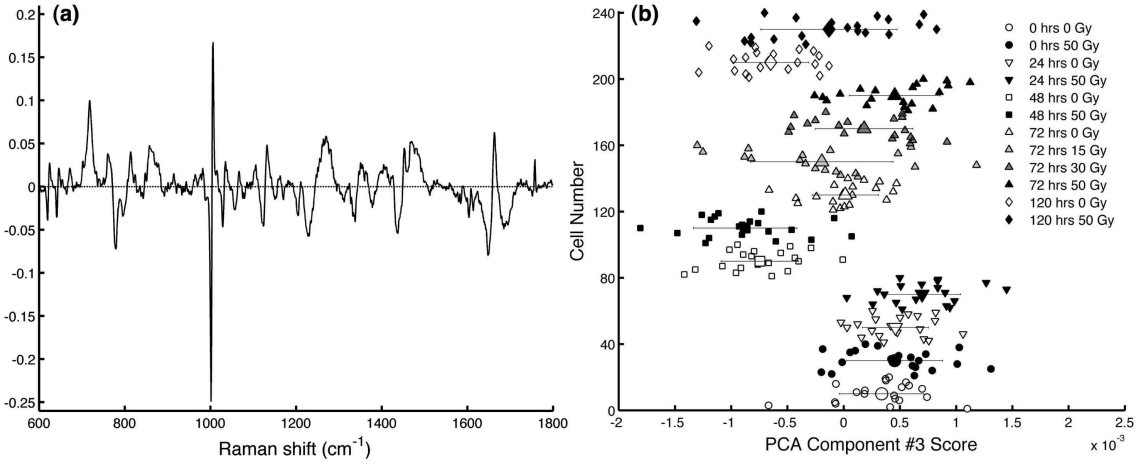


Figure 6.4: (a) Third PCA component (4.7% of total variance) and (b) PCA scores for the LWN spectral window. Different markers categorize all 240 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

In this study, the third PCA component is the last component that displays any trend in the PCA score plots, and each of the remaining 236 PCA components explains less than 4% of the total variance, suggesting that all useful biological information in the data set is contained within the first two PCA components.

HWN window

Recalling figure 6.1b, the HWN window spectrum of a single cell is a superposition of broad features arising primarily from the stretching of CH_2 and CH_3 groups in proteins and lipids. The first PCA component for the HWN window (figure 6.5a) explains 88.9% of the total variance. The positive features in the component arise from proteins and aromatic groups, whereas the negative features are uniquely lipid in origin, thus matching the biomolecular separation observed in the first PCA com-

ponent for the LWN window (figure 6.2a). The HWN window PCA scores for the first PCA component (figure 6.5b) show a nearly identical trend as the LWN window PCA scores (figure 6.2b). As shown in chapter 5 (and the corresponding publication [168]), for DU145 cells this HWN window first PCA component arises from the same inherent cell cycle variability observed in the LWN window spectra, and typically explains 75-90% of the total variance in the data set, as opposed to 40-60% for the LWN window first PCA component.

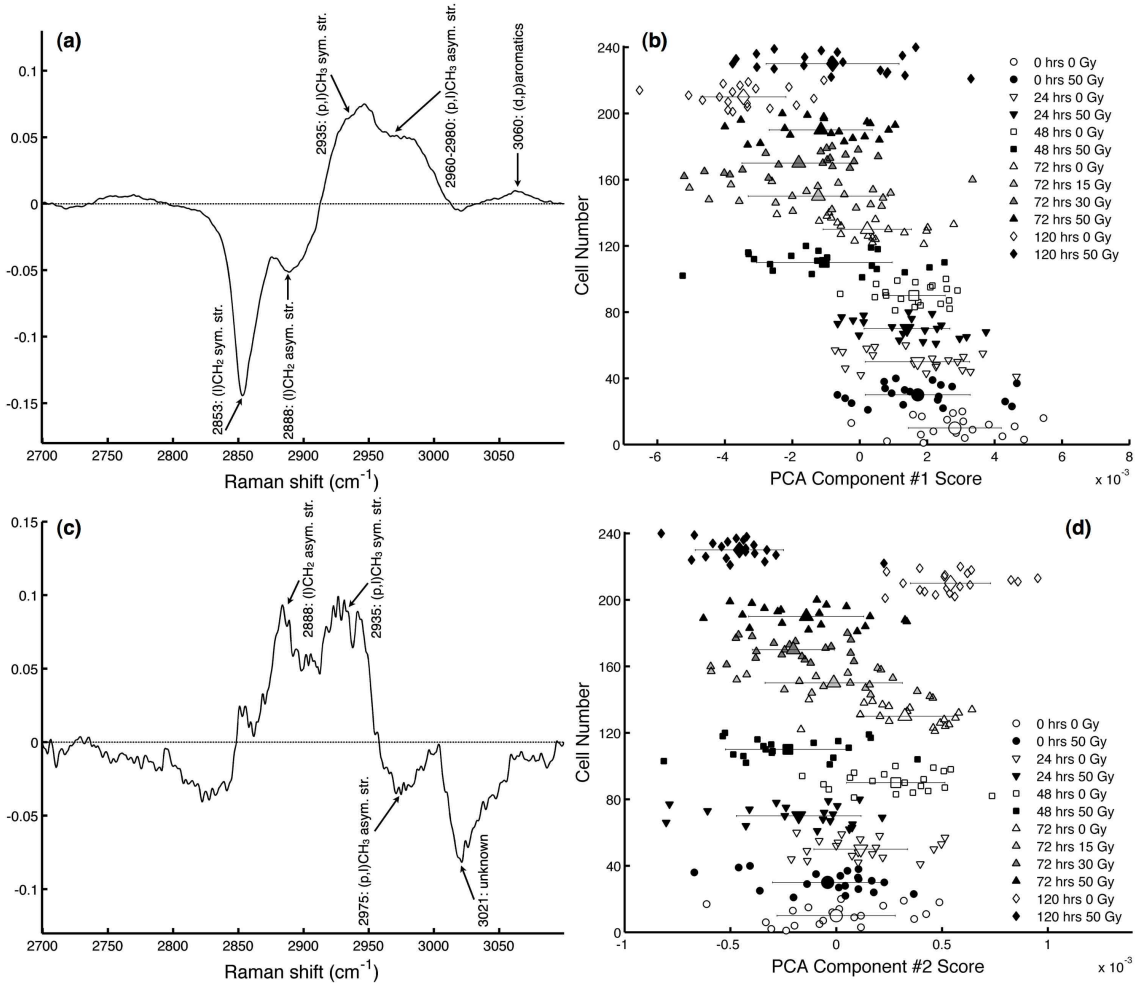


Figure 6.5: First and second PCA component results, for the HWN window. (a) First and (c) second PCA components (88.9% and 2.1% of total variance, respectively), with Raman shifts and molecular origins (figure 6.1b) of identifiable features. Abbreviations - as in figure 6.1. (b) & (d) PCA scores for the (b) first and (d) second PCA components. Different markers categorize all 240 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

The second PCA component for the HWN window (figure 6.5c) explains only 2.1% of the total variance, yet still contains several identifiable features that are consistently reproduced during similarly conducted DU145 irradiation experiments. However, the validity of assigning molecular origins to features in the component is somewhat uncertain due to their broad and somewhat noisy nature. Despite this uncertainty, the HWN window PCA scores (figure 6.5d) display a separation in the average PCA score correlated with both the incubation time post-irradiation and the irradiated dose after 72 hours, which is similar to what is observed for the LWN window PCA scores for the second PCA component (figure 6.3b). However, for the HWN window the separation in PCA scores between 0 Gy and 50 Gy samples is less pronounced at all time points than for the LWN window. Furthermore, the dose dependance of the PCA scores at 72 hours post-irradiation is not reproduced between 30 and 50 Gy, and the incubation time dependence of the PCA scores for the 50 Gy samples is not reproduced between 48 and 72 hours post-irradiation (figure 6.5d).

The remaining PCA components for the HWN window each explain less than 1% of the total variance. The components have no identifiable features and the PCA scores do not display any trends, suggesting that all useful biological information in the data set is contained within the first two PCA components (as was the case for the LWN window data set).

Reproducibility

The results presented above have been reproduced in three separate experiments with DU145 cells (including the experiment discussed below in section 6.3.3). The first and second PCA components for the LWN and HWN window (*e.g.*, figures 6.2a and 6.3a, LWN window) were consistently reproduced. Furthermore, for both spectral windows the PCA scores for the first PCA component followed the expected trend for changes in the cell cycle distribution, and the PCA scores for the second PCA component displayed similar separations in average PCA score, correlated with dose and incubation time post-irradiation, as is observed for the irradiated and unirradiated samples presented here (*e.g.*, figures 6.2b and 6.3b, LWN window). In the first of these experiments, only 80 spectra were obtained for each spectral window. Even though far fewer spectra were obtained compared to the experiment presented above (240 spectra for each window), the features in the first and second PCA components and the trends in the corresponding score plots were successfully reproduced.

6.3.3 Effect of time of irradiation after sub-culturing

Additional experiments have been performed with DU145 cells using the same irradiation and RS acquisition protocols as in the experiment presented in section 6.3.2 (except no data was collected at 48 hours post-irradiation), with the important difference being that cells were irradiated ∼1 day after sub-culturing (at a confluency of 20-30%), instead of ∼3.5 days after sub-culturing (at a confluency of 70-80%). In chapter 5 (and the corresponding publication [168]), it is shown that there is inherent spectral variability observed in the LWN window Raman analysis of single cells from *in vitro* DU145 cell cultures, independent of the existing cell cycle variability, which is correlated with the incubation time of the cell culture up to 3-4 days after sub-culturing. It is also shown that this extra inherent source of variability is not observed if RS analysis began 3-4 days after sub-culturing, or later (recall figure 5.7a). As such, during these additional DU145 irradiation experiments radiation induced spectral differences were observed alongside spectral differences associated with inherent biochemical changes due to incubation time, both of which are correlated with the time of RS analysis after sub-culturing. These competing effects were subsequently observed in the second and third PCA components for the LWN window (figure 6.6). The LWN window results from one of these experiments are presented below.

The first PCA component from this experiment (explaining 38.6% of the variance) and the corresponding PCA scores (not shown) are both very similar to the corresponding results from the experiment in section 6.3.2 (figure 6.2), and similarly describe the inherent cell cycle variability in the data. The second PCA component (figure 6.6a) explains 15.5% of the variance, and contains nearly identical features as the second PCA component from the experiment in section 6.3.2 (figure 6.3a), which explains a comparable amount of the total variance (14.4%) in that data set. The PCA scores for the second PCA component (figure 6.6b) confirm that the spectral differences described by the second PCA component are correlated with both the incubation time post-irradiation and the irradiated dose after 72 hours; however, the trend of increasing separation in the average PCA score with increased incubation time is slightly reversed between 72 and 120 hours, unlike the corresponding PCA score results from the experiment in section 6.3.2 (figure 6.3b). A possible explanation for this discrepancy is provided below. The third PCA component (figure 6.6c) explains 8.1% of the variance, and is not observed in the experiment presented in

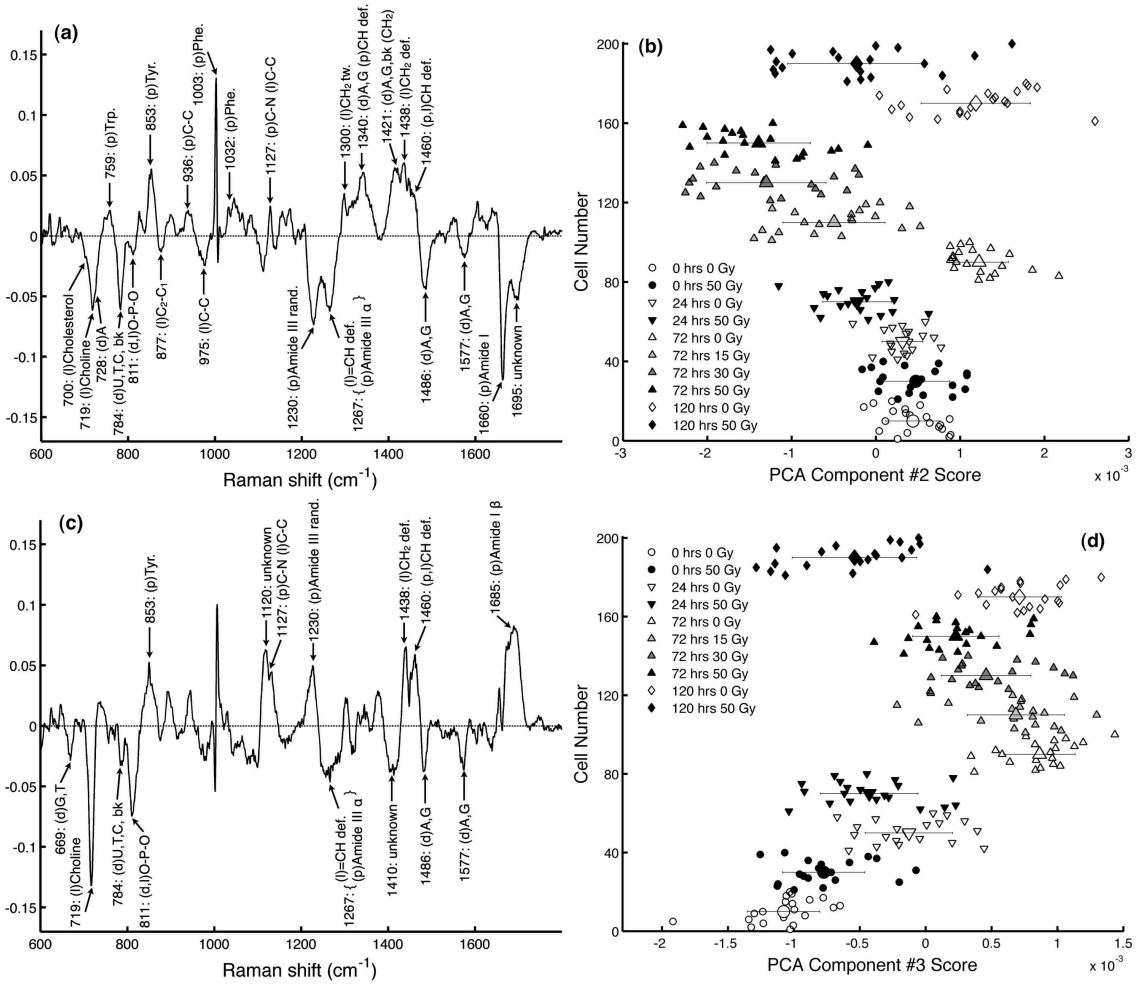


Figure 6.6: Second and third PCA component results, for the LWN window, from a radiation experiment with cells irradiated at 20-30% confluency. (a) Second and (c) third PCA components (15.5% and 8.1% of total variance, respectively), with Raman shifts and molecular origins (figure 6.1a) of identifiable features. Abbreviations - as in figure 6.1. (b) & (d) PCA scores for the (b) second and (d) third PCA components. Different markers categorize all 200 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

section 6.3.2. This new component arises from the additional inherent biochemical variability correlated with the incubation time of DU145 cell cultures in the first 3-4 days after sub-culturing. Almost all of the features in the third PCA component presented here are also observed in PCA components derived from RS studies of unirradiated DU145 cell cultures (figure 5.6a in chapter 5). The average PCA scores for the third PCA component (figure 6.6d) increase from 0 to 72 hours post-irradiation (*i.e.*, 1 to 4 days after sub-culturing), for both unirradiated and irradiated samples. From 72 to 120 hours post-irradiation there is no noticeable change in the average PCA score for the unirradiated samples (consistent with previous observations [168]), and a clear decrease in the average PCA score for the irradiated sample. A possible explanation for this decrease, and for the corresponding aberrant increase in average PCA score for the second PCA component, arises from the presence of several spectral features appearing in both the second and third PCA components. As such, the respective sources of spectral variability caused by irradiation and by incubation time after sub-culturing may not be completely separated into each component. Since the biochemical changes from each effect occur over similar time scales there may be some "mixing" of these two components.[†] Due to the competing biological changes described by the third PCA component, the degree to which the radiation induced spectral variability is uniquely described by the second PCA component in this experiment is unclear. Nevertheless, the similarity of the second PCA component and the corresponding PCA score plot with the results presented in section 6.3.2 lends weight to the conclusion that radiation induced spectral variability is well described by the features in the second PCA component of such irradiation RS experiments on DU145 cell cultures.

6.4 Discussion

6.4.1 First PCA component: Cell cycle variability

In the experiments presented in this chapter, the largest source of spectral variability between single live cells, explained by the first PCA component (figure 6.2a) as varying

[†]The HWN window results for this experiment (not shown) corroborate this hypothesis. The inherent spectral variability correlated with the incubation time of DU145 cell cultures in the first 3-4 days is much reduced in the HWN window as compared to the LWN window, as shown in chapter 5. As such, in the irradiation experiment described here the HWN window PCA scores for the second PCA component do not show the same decrease in average score separation between 72 and 120 hours that is observed for the LWN window PCA scores (figure 6.6b).

concentrations of proteins and nucleic acids relative to lipids, arises from inherent differences in cell cycle between cells within a given culture and between cultures. Therefore, irradiation of DU145 cell cultures does not induce more spectral differences in the Raman spectra of surviving cells, up to 5 days post-irradiation, than what is already explained by cell cycle biochemical differences. However, the delivery of high doses of radiation may significantly alter the cell cycle distribution of a culture over time [31], irrespective of any uniquely radiation induced biochemical changes. Hence, the cell cycle distributions of all cultures collected in this work were measured with flow cytometry (figure 6.7) to verify that the observed cell cycle distributions were consistent with the spectral variability described by the first PCA component (section 6.3.2, figure 6.2).

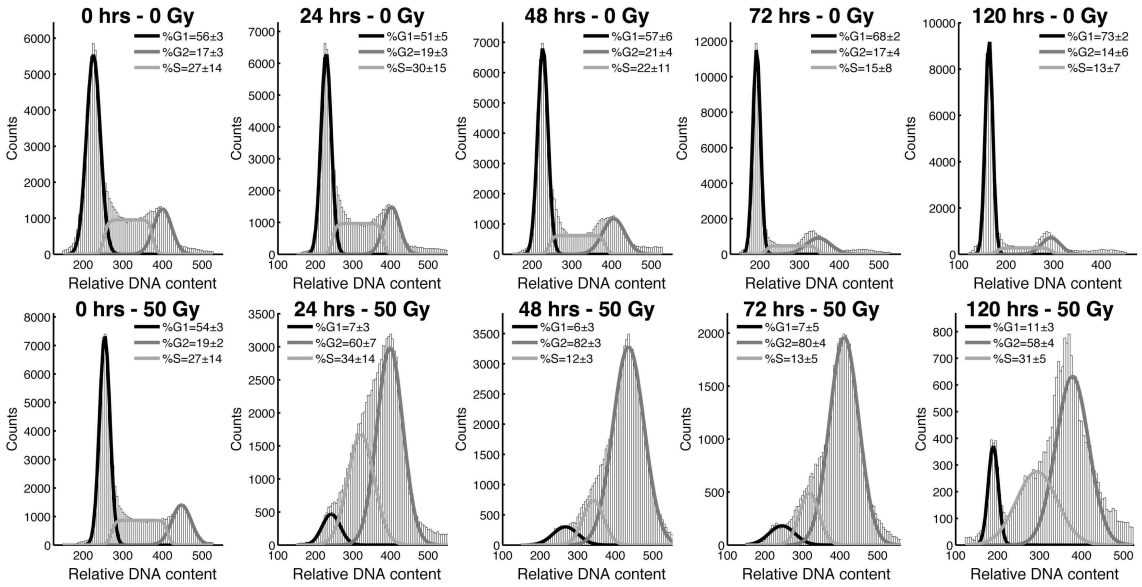


Figure 6.7: Flow cytometry analysis of unirradiated (0 Gy, top row) and irradiated (50 Gy, bottom row) cell cycle distributions. The number of hours shown is the incubation time of the culture after irradiation.

The cell cycle distributions of the unirradiated cell cultures (top row, figure 6.7) remain relatively unperturbed until ∼72 hours where there is a significant increase in the fraction of G1 phase cells and a corresponding decrease in the fractions of S and G2 phase cells, a trend that continues for the 120 hours cell culture. This observation is expected, since by 72 hours the unirradiated cell cultures had reached confluency and had begun to progress into a less proliferative state with an increased fraction of quiescent (nonproliferating G1 phase, or 'G0' phase) cells. The PCA scores for the first PCA component for the unirradiated cell cultures are well correlated with

these changes in the cell cycle distributions, as the PCA scores (figure 6.2b) indicate that the average protein and nucleic acid content relative to lipid content remains relatively constant until decreasing continually from 48 to 72 to 120 hours. This trend is consistent with previous observations (chapter 5 and the corresponding publication [168]) and with known cell cycle related biochemical changes in nonproliferative and actively cycling cells [17, 18, 83].

At all times after irradiation the cell cycle distributions of the irradiated cell cultures (bottom row, figure 6.7) are significantly altered from a cell cycle distribution characteristic of an asynchronously dividing proliferative culture (*i.e.*, the two left-most panels in figure 6.7). However, from 0 to 24 hours post-irradiation the PCA scores for the irradiated cultures are similar to the PCA scores for the unirradiated cultures, whereas the 50 Gy irradiated cultures at 48, 72 and 120 hours have lower PCA scores with average values comparable to the average PCA score for the 120 hours unirradiated culture (figure 6.2b). Recall that flow cytometry is used here to measure the relative DNA content of cells. As such, at 24 hours after irradiation the surviving cells have begun to be arrested in G2 phase (double the DNA content of a G1 phase cell). The G2 arrest is most pronounced for the 48 and 72 hours samples, and for each of the 24, 48 and 72 hours samples a distinct G1 peak is not observed. At 120 hours post-irradiation a G1 peak is observed, possibly due to some fraction of G2 arrested cells having successfully repaired enough radiation damage to progress through a mitotic division. The type of cell cycle redistribution observed here in response to single high doses of photons is consistent with results reported from a previous radiation study with DU145 cells, where the measured fraction of G2 cells approached 50% at 72 hours after irradiation to 20 Gy [179].

In chapter 5 (and the corresponding publication [168]) it is shown that between late G1, S and G2 phase cells, the differences in nucleic acid and protein content relative to total lipid content are very small, as measured with RS (recall figure 5.12). The largest cell cycle related spectral differences arise from early G1 cells, which are known to have the lowest levels of RNA and protein among actively proliferating cells [17, 18]. Even more pronounced spectral differences are observed for nonproliferative cells [83, 168], which are known to have an even lower RNA and protein content than early G1 cells [17, 18]. Since the irradiated cell cultures in this work become nonproliferative as a result of radiation damage (figure 6.7), the observed decrease in the average nucleic acid and protein content relative to lipid content of the irradiated cell cultures from 48 to 120 hours post-irradiation (figure 6.2b) may be explained

by radiation induced nonproliferation causing similar cell cycle related biochemical changes as those observed in unirradiated nonproliferative cultures. However, the possible biological mechanisms responsible for such changes are unknown.

6.4.2 Second PCA component: Radiation induced effects

The second PCA component in this study (figure 6.3a) describes RS spectral differences that are independent from the existing cell cycle related spectral differences described by the first PCA component. The PCA scores (figure 6.3b) show that the magnitude of the spectral differences increases with the incubation time post-irradiation and with the irradiated dose after a fixed incubation period. The separation in PCA score distributions between the 50 Gy irradiated and the unirradiated cultures is not significant immediately after irradiation ($p = 0.15$ at 0 hrs), but is statistically significant by 24 hours ($p < .0001$), and at all times afterwards. At 72 hours post-irradiation, increasing the dose from 0 to 15 to 30 to 50 Gy results in corresponding statistically significant decreases in the PCA scores ($p < .0001$ for any pair of samples). The additional experiments described in sections 6.3.2 and 6.3.3 confirm that the spectral differences described by the second PCA component (figure 6.3b, section 6.3.2) are indeed radiation induced effects.

These results show that over time, DU145 cells in culture irradiated to single high doses display increased levels (negative features in second PCA component for the LWN window, figure 6.3a) of choline, O-P-O groups in lipids and RNA, C-C and =CH groups in lipids, random coil and α-helix amide groups, nucleic acid bases and DNA, as well as decreased levels (positive features in second PCA component for the LWN window, figure 6.3a) of aromatic amino acids, C-C, C-N and CH groups in proteins, CH_2 groups in nucleic acids, and CH_2 and CH groups in lipids. No wavenumber shifts or changes in the spectral line shapes of Raman features are observed post-irradiation (*e.g.*, figure 6.3c), indicating that the spectral changes observed here are the result of biological responses to radiation, rather than chemical changes (further discussed below). The Raman spectral changes corresponding to these biochemical changes are summarized in table 6.1.

It should be emphasized that these biochemical changes occur independent from any biochemical changes related to cell cycle progression described by the first PCA component, even though the cell cycle progression of the cultures are altered by irradiation (section 6.4.1, figure 6.7). The spectral differences described by the second

Table 6.1: Raman spectral changes correlated with dose (0 to 50 Gy) and post-irradiation incubation time (0 to 120 hours), observed in irradiated DU145 cells. ↑ / ↓ = increased / decreased Raman intensity with increased dose and incubation time (see figures 6.3a and 6.3b). Abbreviations - as in figure 6.1.

	Raman Shift (cm^{-1})	**Molecular Assignment**	**Raman Intensity**	
DNA/RNA	784 {	U, C, T ring br, O-P-O str bk }	↑	
	811	O-P-O str RNA	↑	
	1421	CH_2 bk		↓
	1486	A, G	↑	
	1577	A, G	↑	
Proteins	759	Trp ring br		↓
	853	Tyr ring br		↓
	936	C-C str bk, α-helix		↓
	1003	Phe ring br		↓
	1032	C-H Phe		↓
	1127	C-N str		↓
	1230	Amide III rand coil	↑	
	1267	Amide III α-helix	↑	
	1340 & 1460	CH def		↓
	1660 {	Amide I α-helix Amide I rand coil }	↑	
Lipids	719	Choline	↑	
	811	O-P-O str	↑	
	877	Acyl C_2-C_1 str	↑	
	975	Head C-C str	↑	
	1127	Chain C-C str		↓
	1267	=CH def	↑	
	1300	CH_2 tw		↓
	1438	CH_2 def		↓
	1460	CH def		↓

PCA component are uniquely radiation induced biochemical changes in single DU145 cells. Interestingly, the biochemical nature of the observed radiation response is the same for all surviving cells analyzed from 24 to 120 hours post-irradiation, and it is the magnitude of the observed biochemical changes that increases with the post-irradiation incubation time and the irradiated dose. It should also be emphasized that since there is no significant separation in the PCA score distributions between the 50 Gy irradiated and the unirradiated cultures immediately after irradiation (figure 6.3b, '0 hrs 0 Gy' vs. '0 hrs 50 Gy'), the current acquisition and analysis methods are only sensitive to biochemical changes describing the response of single cells to irradiation over time, and are not sensitive to the immediate physical or chemical changes caused by direct or indirect radiation damage.

Several studies have investigated radiation induced changes in cultured cells with other methods, such as Fourier transform infrared microscopy (FTIRM) and magnetic resonance spectroscopy (MRS). Direct comparisons of the FTIRM studies with the RS results presented here are difficult due to the different molecular sensitivities between FTIR and RS, the different spectral processing methods used (most notably the method of normalization), the different time intervals between irradiation and analysis, and the presence or absence of cell cycle related spectral changes. However, three previous FTIRM studies [180–182] have reported radiation induced spectral changes in HaCaT cells (non-tumourogenic immortalized keratinocyte cell line), many of which corroborate the results presented here (table 6.1). All three studies reported changes in conformational protein structure, decreases in tyrosine signal, and decreases in CH_2 lipid signal after irradiation, each of which support the spectral changes observed here (table 6.1). Furthermore, the most recent FTIRM study [182] also observed spectral features changing with increased dose and incubation time, up to 96 hours post-irradiation, qualitatively matching the trends observed in our study. One MRS study [133] examined changes in metabolites in HeLa cells after similar doses (40 Gy) and post-irradiation incubation times (24 to 72 hours) as used in the current work. This MRS study identified an increase in choline signal that correlated with increased incubation time post-irradiation, which qualitatively corroborates the spectral changes in choline intensity observed here (table 6.1).

Although previous studies have reported on the use of RS to detect radiation induced biochemical changes in biological systems, it is difficult to corroborate the results of previous Raman studies with the results presented here because the biological systems undergoing irradiation are completely different. Raman studies have been

performed on isolated DNA [85, 86], lipid membranes [61, 88–91], proteins [61, 89, 90] and tissues [92, 93]. In this work, the biochemical changes observed reflect the response of single cells to irradiation over time, rather than chemical changes caused by direct radiation damage. In effect, this study observes the biological radiation response of surviving cells,[‡] as opposed to a physical or chemical radiation response. Furthermore, this study observes radiation induced spectral changes independent of any spectral changes arising from differences in cell cycle progression, which are observed concurrently with any radiation induced effects. Therefore, Raman studies such as those mentioned above are generally expected to yield different results than a Raman study performed with cells in culture. This study is the first time vibrational spectroscopy has been applied to a living biological system to observe radiation induced biochemical changes that are independent of biochemical changes arising from concurrent cell cycle or cell death related processes.

6.4.3 LWN vs. HWN spectral windows

The HWN window offers several technical advantages over the LWN window, such as reduced background fluorescence, reduced acquisition time to cover the spectral window, and simplified baseline estimation. However, there is a considerably higher number of detectable molecular groups in the LWN window, which may translate into an increased diagnostic benefit for measuring radiation induced effects. For the experiment presented in section 6.3.2, there is an increase in the PCA score separations between unirradiated and irradiated cultures for the LWN window (figure 6.3b) as compared to the HWN window (figure 6.5d). Furthermore, the observed trend of increased score separation with incubation time and dose is not consistent for all samples in the HWN window, as is observed for the LWN window. Specifically, the LWN window trend is reproduced in the HWN window data for all cultures except the '72 hrs 50 Gy' culture (figure 6.5d). The exact cause of this specific discrepancy is unclear, but is likely due to the very small number of molecular groups available for detection in the HWN window compared to the LWN window, thus decreasing the sensitivity of detecting subtle molecular changes with RS and PCA. This hypothesis is supported by the very low percent variance explained by the second PCA component for the HWN window of 2.1%, as compared to 14.4% for the LWN window. A

[‡]It should be noted that in general the surviving cells analyzed in this work are not viable clonogenic cells, but are rather functioning metabolizing cells that would likely have undergone cell death if allowed to incubate for longer periods of time.

final disadvantage of the HWN window is that the biomolecules responsible for the radiation induced effects, as given by the features in the second PCA component, cannot all be assigned with confidence. In summary, for the application of RS to study radiation induced effects in single cells, the LWN window provides increased diagnostic utility and more biochemically useful information.

6.5 Conclusion

This chapter presents the first use of RS to observe radiation induced biochemical changes in single human cells cultured *in vitro*. The sensitivity of RS to different types of biomolecules in a single acquisition allows the noninvasive and nondestructive observation of radiation induced effects in single living cells. PCA is shown to effectively discriminate between radiation induced spectral changes and other inherent sources of spectral variability, such as those arising from cell cycle. The radiation induced spectral changes in irradiated DU145 cells are assigned to various nucleic acids, lipids, amino acids and conformational protein structures (table 6.1), and are found to correlate with both the irradiated dose (0 to 50 Gy) and the post-irradiation incubation time (0 to 120 hours). This successful demonstration of the sensitivity of RS to radiation induced changes in single cells makes RS an attractive modality for further radiobiological studies of living biological systems.

Chapter 7

Results & Discussion IV: Biochemical signatures of radiation response in lung, breast and prostate tumour cells

7.1 Introduction

The previous chapter demonstrates the sensitivity of RS to radiation induced changes in single DU145 cells, proving that RS is indeed an attractive modality for further radiobiological studies. As such, the radiation response of different cell lines is investigated here with RS. Of immediate interest in radiobiology is the biochemical reason for differences in intrinsic radiosensitivity across different cell lines. The purpose of the studies and analyses presented in this chapter is to investigate any correlation between the biochemical radiation response observed with RS and the intrinsic radiosensitivity of a variety of tumour cell lines.

This chapter presents the results from two studies. The first study, presented in sections 7.2 to 7.4, uses previous techniques and protocols (chapter 6) to investigate the *in vitro* radiation response of a variety of human tumour cell lines that vary by tissue of origin, p53 gene status and intrinsic radiosensitivity (quantified by the surviving fraction after 2 Gy (SF_2)). The second study, presented in sections 7.5 and 7.6, investigates the radiation response of a mouse prostate tumour cell line (TRAMP-C2), investigated both *in vitro* (using previous techniques and protocols) and *in vivo*, by irradiation of solid tumours grown in mice.

The first study extends the previous methods used for DU145 cells (chapter 6) to a panel of seven human tumour cell lines derived from prostate (DU145, PC3, LNCaP and PacMet), breast (MDA-MB-231 and MCF7) and lung (H460). Within the tested cell lines, RS acquisition in the LWN spectral window and application of PCA is shown to detect biochemical radiation response signatures that segregate according to intrinsic radiosensitivity and p53 status. For all cell lines, PCA effectively distinguishes radiation induced biochemical changes from any biochemical changes arising from cell cycle differences or other factors. The observed radiation response signatures indicate changes in the cellular concentration of aromatic amino acids (tyrosine, tryptophan and phenylalanine), conformational protein structures (α-helices, β-sheets and random coils), and certain nucleic acid and lipid functional groups. These radiation induced biochemical changes are detected within the first 1-3 days post-irradiation. Many of the observed biochemical changes are consistent with known cellular response mechanisms to radiation exposure (*e.g.*, synthesis and degradation of structured proteins), whereas others may be novel discoveries. The observed relationship between the RS radiation response signatures and intrinsic radiosensitivity may indicate the detection of biochemical mechanisms of radiation resistance or sensitivity with RS. Some candidate processes for such mechanisms are discussed (section 7.4.3).

The second study presents a preliminary comparative investigation of *in vitro* versus *in vivo* radiation response as measured with RS, using the mouse prostate tumour cell line TRAMP-C2. This study is a pilot experiment performed in collaboration with members of the Radiation Cancer and Immunotherapy Theme (RCIT) group at the BC Cancer Agency's Vancouver Island Centre, who were independently investigating the interaction between radiation and the immune system using TRAMP-C2 tumours grown and irradiated in immunocompetent mice (*i.e.*, mice with intact immune responses). As such, the protocols for tumour growth and irradiation (section 7.5.3), and the cell line used (TRAMP-C2), were determined by the RCIT group and the University of Victoria's Animal Care Committee. In this work the *in vitro* radiation response of the TRAMP-C2 cell line is investigated using the same methods and protocols used for the other *in vitro* irradiation experiments presented in the first study. The *in vivo* radiation response is investigated by implanting tumour cells into the subcutaneous flank of mice, monitoring solid tumour growth, irradiating the solid tumours, and then extracting both irradiated and unirradiated tumour cells for RS analysis post-irradiation. A weak radiation response signature is observed *in vitro*, which correlates with many of the response signatures obtained from the hu-

man tumour cell lines. A similar yet strengthened response signature is observed *in vivo*. However, the TRAMP-C2 spectral variability assigned as a radiation induced response, observed both *in vitro* and *in vivo*, is also observed from TRAMP-C2 data sets comprised of only unirradiated tumour cells. Possible explanations for this are discussed, with respect to the radiation responses observed *in vitro* for the human tumour cell lines. Difficulties associated with the *in vivo* pilot study, and possible resolutions, are briefly discussed (section 7.6.4).

7.2 Materials & Methods: Study #1

The cell lines used in this study vary according to p53 status and average reported radiosensitivity (see section 7.3.5), given by average SF_2. The SF_2 values quoted here are calculated from literature data from 4 or 5 different laboratories per cell line [26, 55, 183–195] whose methods closely approximate accepted protocols for clonogenic survival assays [27]. However, no radiosensitivity parameters for the PacMet cell line have been reported to date.

Prior to RS analysis, the cell cycle distribution and viability of each culture was determined with flow cytometry, as described previously. For all experiments in this work, the fraction of live cells in the harvested samples was between 75% and 99%. Any cell spectrum resembling that of a dead cell [59] was rejected during RS data acquisition. No correlations were found between the viability of harvested cultures and the RS data collected, confirming that the cell spectra acquired for processing and PCA analysis were indeed obtained from live cells.

In this study, cultures irradiated to 50 Gy were harvested for RS analysis at 0, 24, 48 and 72 hours post-irradiation, with one unirradiated culture harvested and analyzed at the same time. As before (chapter 6), cultures irradiated to 15 and 30 Gy were harvested and analyzed at 72 hours post-irradiation. Data from 120 hours post-irradiation was only acquired if the viability of both cultures exceeded 75%; in this study the LNCaP, MCF7 and PacMet cell lines did not satisfy this criteria. As such, for all cell lines in this study only the RS data collected from 0 to 72 hours post-irradiation is included, in order to facilitate comparisons between data sets.

RS acquisition and data processing was performed as described previously. Only the LWN spectral window data is presented here, as the analysis of the HWN spectral window data demonstrated an insufficient sensitivity to biochemical radiation responses in that region. No sources of variability arising from the quartz substrate

were observed in this study; as such, no filtering was performed with PCA.

7.3 Results: Study #1

7.3.1 Unirradiated cell spectra

The Raman spectra of single unirradiated DU145 cells, from 600-1800 cm^{-1} (figure 7.1a), and the molecular assignments for the spectral features, are presented and discussed in previous chapters (and the corresponding publications [168, 178]). For comparisons with the cell lines used in this study, the averaged Raman spectrum from 20 unirradiated cells for all seven cell lines, harvested and analyzed immediately after the time of irradiation, is presented in figure 7.1b. All spectral features observed for DU145 cells (figure 7.1a) are also observed for the other cell lines. There are subtle differences in the relative intensity of spectral features between cell lines, arising from inherent differences in biomolecular compositions between tumour cell lines [79] and slightly different cell cycle distributions between cultures at time of irradiation [168].

7.3.2 Cell cycle spectral variability

The previous studies in chapter 6 (and the corresponding publication [178]), using the prostate cell line DU145, demonstrate that Raman spectral variability arising from inherent cell cycle differences between cells is identified by the first PCA component of a RS data set obtained from both unirradiated and irradiated DU145 cells. In this study with multiple cell lines, the cell cycle PCA component observed for DU145 cells (figure 7.1c) is consistently reproduced as either the first or second PCA component of all data sets (figure 7.1d) and accounts for 10-50% (figure 7.1d) of the total variance in the data, depending on the cell line. Recall that this cell cycle variability arises from decreased concentrations of protein and nucleic acids (positive features in the PCA components, figures 7.1c and 7.1d) relative to lipids (negative features in the PCA components, figures 7.1c and 7.1d) in early G1 (or G0) phase cells as compared to late G1, S and G2 phase cells [168].

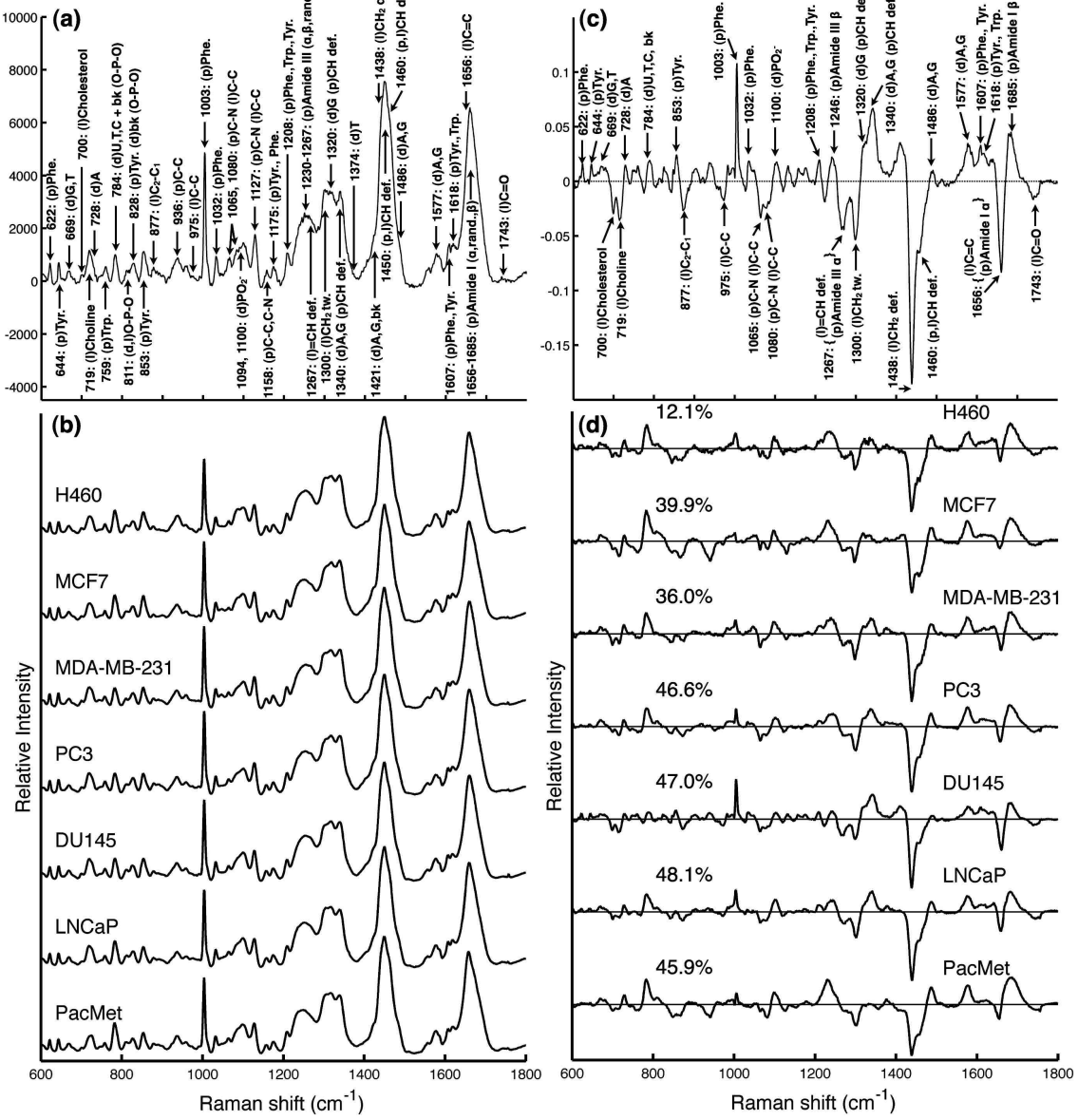


Figure 7.1: (a) Sample Raman spectrum of a single unirradiated DU145 cell. (b) Average spectra from 20 unirradiated cells for the seven cell lines used in this study. (c) Cell cycle PCA component for the DU145 data set (200 cells). (d) Cell cycle PCA components for all seven cell lines (200 cells each), with percent variance explained by each component. The Raman shift and molecular origin of identifiable features are provided in (a) and (c) [59, 75, 92, 118–121]. Abbreviations - p: protein, l: lipid, d: DNA/RNA, A: adenine, T: thymine, G: guanine, C: cytosine, U: uracil, Phe: phenylalanine, Tyr: tyrosine, Trp: tryptophan, bk: backbone, def: deformation, tw: twist, sym: symmetric, asym: asymmetric, str: stretch.

7.3.3 RS radiation response signatures I: Categories R1, R2 and R3

PCA components corresponding to radiation induced biochemical changes (independent from cell cycle related changes that are described by the cell cycle PCA components, figure 7.1d) are identified by statistically significant ($p < 0.05$) changes in the corresponding PCA score distributions for irradiated samples, as compared to unirradiated samples, in the first 24 to 72 hours after irradiation. In this study, one radiation induced PCA component is identified for each cell line (figure 7.2).

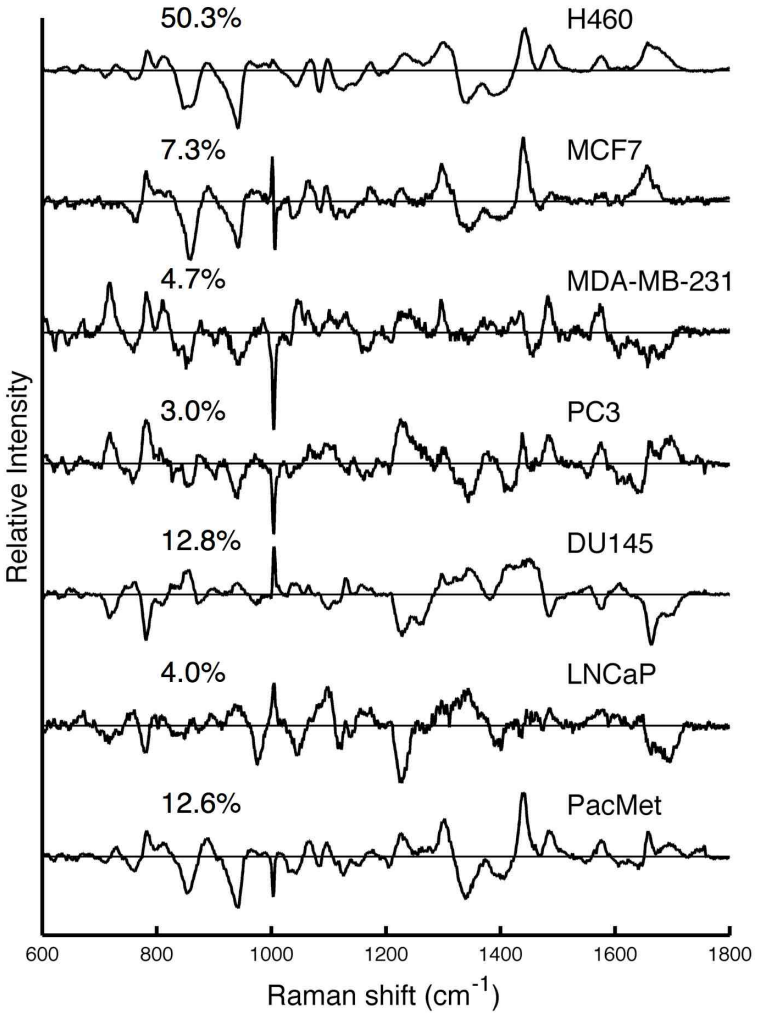


Figure 7.2: Radiation induced PCA components for all 7 cell lines, ordered by tissue of origin, with percent variance explained by each component. The Raman shifts and molecular origins of identifiable features are provided in figures 7.3a & 7.4a, 7.3c & 7.4c and 7.3e & 7.4e.

Correlation analysis between the radiation induced PCA components (table 7.1) indicates that these RS "radiation response signatures" cluster into 3 distinct categories. The radiation response signatures for the H460, MCF7 and PacMet cell lines demonstrate the highest correlations (H460 and MCF7: $r = 0.87$, H460 and PacMet: $r = 0.88$, MFC7 and PacMet: $r = 0.83$) and are therefore assigned to category "R1". The highest correlation with the MDA-MB-231 cell line response signature comes from the PC3 cell line ($r = 0.59$), and these two cell lines are therefore assigned to category "R2". The DU145 and LNCaP response signatures are positively correlated only with each other ($r = 0.51$) and are therefore assigned to category "R3". There is also consistent correlation between the PCA components in groups R1 and R2 ($0.18 < r < 0.67$), and consistent anti-correlation between the two PCA components in group R3 and the other five radiation induced components ($-0.13 < r < -0.70$). All correlations values between PCA components in table 7.1 are statistically different than zero ($p < 0.05$).

Table 7.1: Correlation r-values between radiation induced PCA components (figure 7.2) for each human tumour cell line (1 = perfect correlation, -1 = perfect anti-correlation, 0 = no correlation). All correlation values are statistically different than zero ($p < 0.05$).

Cell Line	**H460**	**MCF7**	**MDA-MB-231**	**PC3**	**DU145**	**LNCaP**	**PacMet**
H460	-	0.87	0.18	0.50	-0.36	-0.17	0.88
MCF7	0.87	-	0.20	0.42	-0.27	-0.13	0.83
MDA-MB-231	0.18	0.20	-	0.59	-0.28	-0.18	0.38
PC3	0.50	0.42	0.59	-	-0.70	-0.48	0.67
DU145	-0.36	-0.27	-0.28	-0.70	-	0.51	-0.32
LNCaP	-0.17	-0.13	-0.18	-0.48	0.51	-	-0.22
PacMet	0.88	0.83	0.38	0.67	-0.32	-0.22	-

7.3.4 RS radiation response signatures II: Radiation induced changes in biomolecules across categories R1, R2 and R3

The biomolecules responsible for the observed radiation induced PCA components (figure 7.2) are identified by the positive and negative features in the PCA components. Molecular assignments are provided in figure 7.3 for one cell line from each RS category, using cell lines H460 (R1, figure 7.3a), MDA-MB-231 (R2, figure 7.3c) and DU145 (R3, figure 7.3e). The radiation induced changes in these biomolecules are given by the changes in the corresponding PCA score distributions for irradiated

cultures in the first 1-3 days post-irradiation, relative to the unirradiated cultures (figures 7.3b, 7.3d and 7.3f). For the PCA score plots shown in figures 7.3b, 7.3d and 7.3f, all irradiated samples demonstrated statistically significant ($p < 0.05$) decreases in their PCA scores from 24 to 72 hours post-irradiation, when compared with unirradiated controls.

The radiation induced PCA components and PCA scores for the MCF7 (R1), PC3 (R2) and LNCaP (R3) cell lines are shown in figure 7.4. The PCA scores show similar changes in their distributions with time and dose as their RS category counterparts (figure 7.3), although with reduced separations and some losses of significance ($p > 0.05$) at certain time points, for certain cell lines. The MCF7 (R1) PCA scores (figure 7.4b) for irradiated cultures are significantly decreased from unirradiated cultures at 24 and 48 hours post-irradiation, as are the 15 and 30 Gy cultures at 72 hours post-irradiation. The 72 hours 50 Gy culture is significantly decreased from the 72 hours 0 Gy culture, but has higher scores than both the 72 hours 15 Gy and the 72 hours 30 Gy cultures. The PC3 (R2) scores (figure 7.4d) for the 50 Gy irradiated cultures are significantly decreased from the unirradiated cultures at 24, 48 and 72 hours post-irradiation, but the 15 Gy and 30 Gy cultures are not. The LNCaP (R3) PCA scores (figure 7.4f) for all irradiated cultures are significantly decreased from the unirradiated cultures from 24 to 72 hours post-irradiation. However, there are no significant differences in the score distributions between the 15, 30 and 50 Gy cultures at 72 hours post-irradiation, unlike as was observed for the DU145 (R3) score distributions (figure 7.3f). The PCA scores for the radiation induced PCA component for the PacMet (R1) cell line (not shown) demonstrate significant decreases in all irradiated cultures at 48 and 72 hours post-irradiation, but not at 24 hours post-irradiation.

For the PCA components from RS categories R1 and R2 (*e.g.*, figures 7.3a, 7.3c, 7.4a and 7.4c), the observed decrease in the PCA scores for irradiated cells (*e.g.*, figures 7.3b, 7.3d, 7.4b and 7.3d) corresponds with increases in amino acids, α-helix protein structure and CH groups (common negative features in R1 and R2 PCA components), and with decreases in nucleic acids, CH_2 groups, and β-sheet and random coil protein structures (common positive features in R1 and R2 PCA components). The presence of many common biomolecules in the PCA components between groups R1 and R2 is reflected by the positive correlation between PCA components in these groups (table 7.1, $0.18 < r < 0.67$). For the RS category R3 components (figures 7.3e and 7.4e), the observed decrease in the PCA scores for irradiated cells (figure 7.3f)

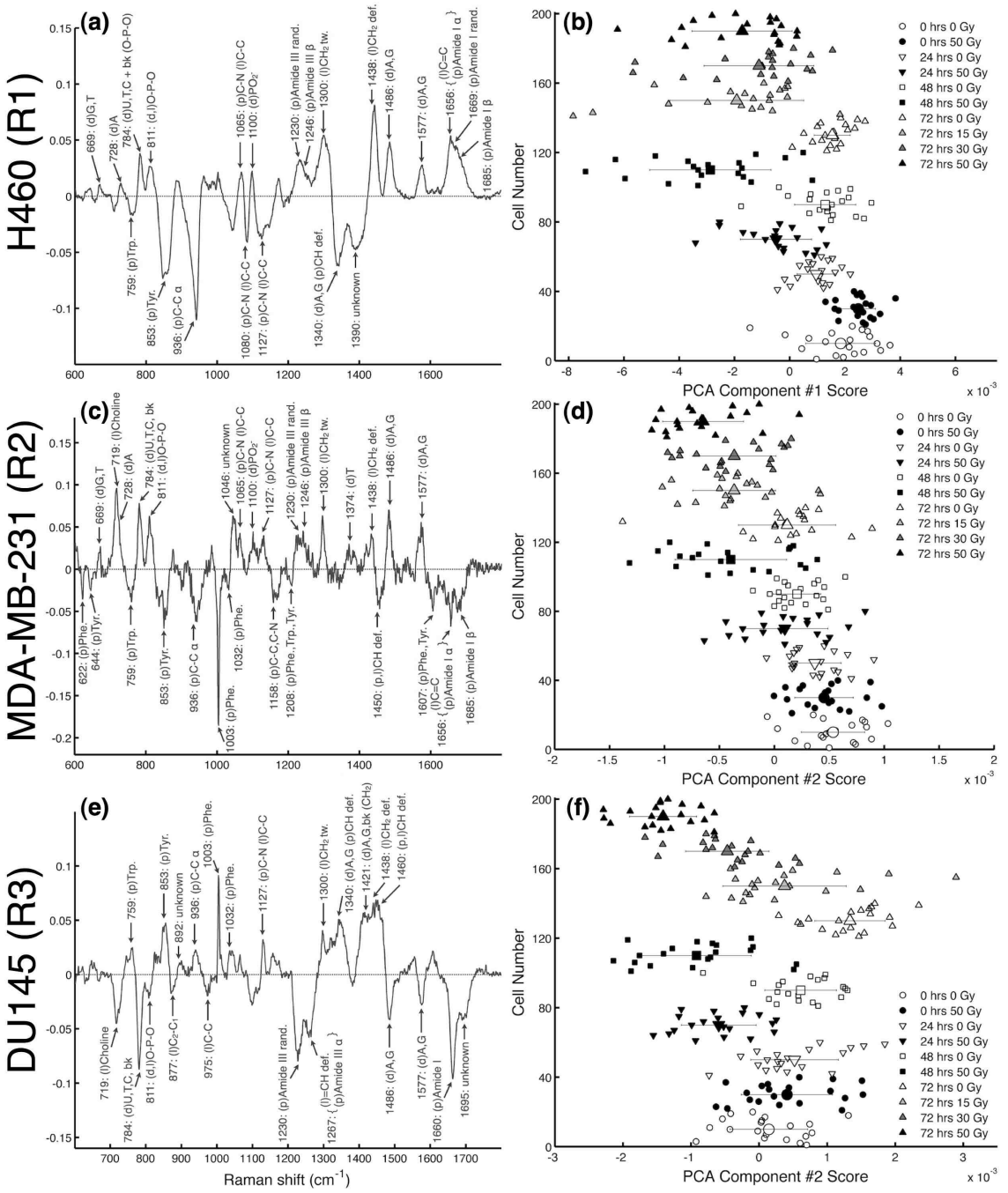


Figure 7.3: (a, c, e) Radiation induced PCA components: (a) H460, (c) MDA-MB-231 and (e) DU145 cell lines. Raman shifts and molecular origins of identifiable features are provided. (b, d, f) PCA scores for the (b) H460, (d) MDA-MB-231 and (f) DU145 radiation induced PCA components. Different markers categorize all 200 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data. Abbreviations - as in figure 7.1.

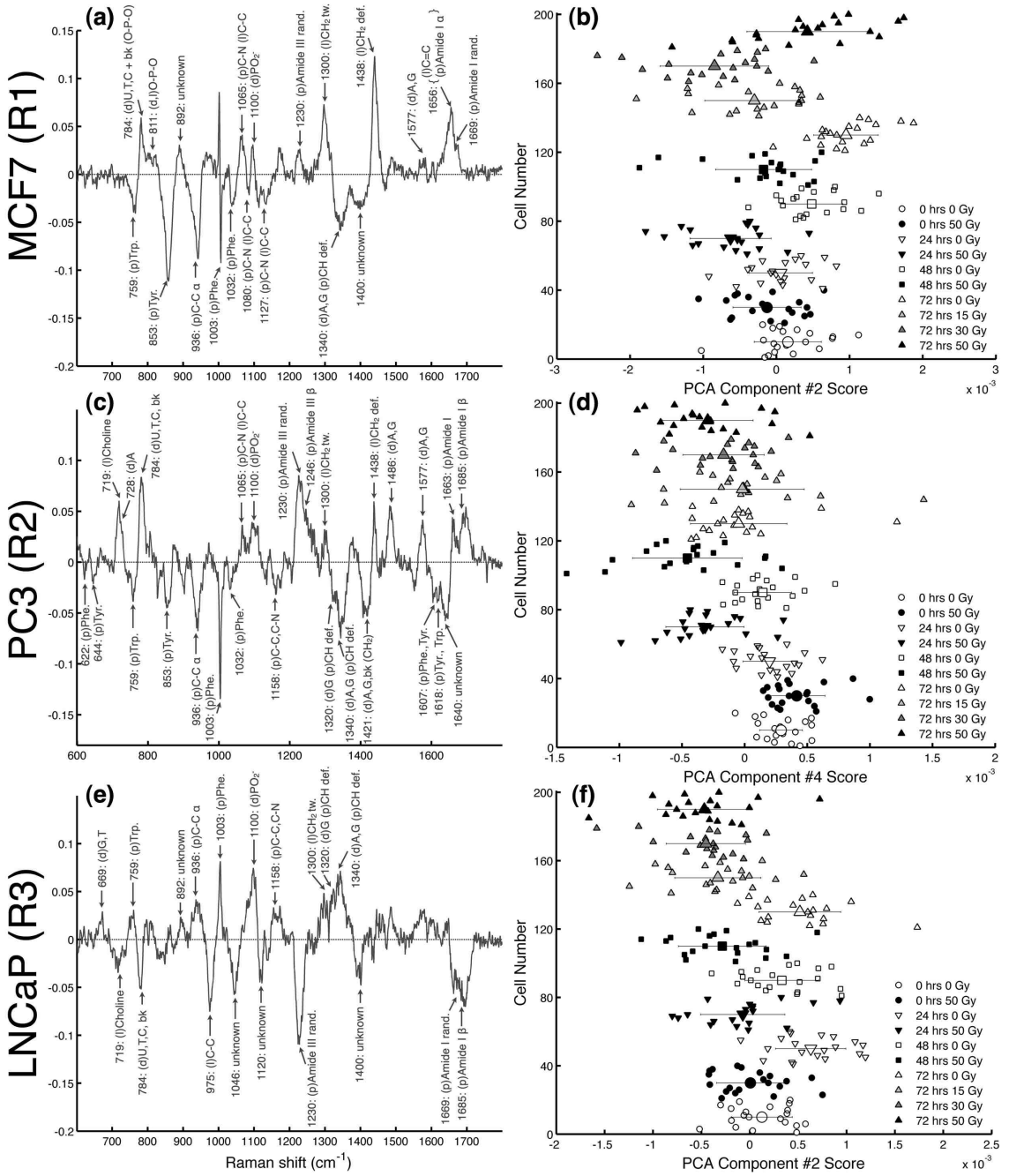


Figure 7.4: (a, c, e) Radiation induced PCA components: (a) MCF7, (c) PC3 and (e) LNCaP cell lines. Raman shifts and molecular origins of identifiable features are provided. (b, d, f) PCA scores for the (b) MCF7, (d) PC3 and (f) LNCaP radiation induced PCA components. Different markers categorize all 200 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data. Abbreviations - as in figure 7.1.

corresponds with decreases in amino acids, α-helix protein structure, and both CH_2 and CH groups (common positive features in the R3 PCA components), and with increases in nucleic acids, β-sheet and random coil protein structures, and choline (common negative features in the R3 PCA components). The presence of many common biomolecules in the RS category R3 components that show opposite radiation induced changes compared to the R1 and R2 components is reflected by the negative correlation values between PCA components in group R3 and groups R1 and R2 (table 7.1, $-0.13 < r < -0.70$).

7.3.5 RS radiation response associations with cell cycle arrest, p53 and radiosensitivity

The RS radiation response categories R1, R2 and R3 segregate according to the known radiosensitivity of the cell lines (table 7.2). The R1 and R2 category cell lines, for which SF_2 values are available, are known to be radiation resistant ($SF_2 > 0.6$). The R3 cell lines are comparatively radiation sensitive ($SF_2 < 0.5$). Furthermore, the R1 cell lines contain a wild-type (wt) p53 gene, whereas the R2 cells contain a mutant (mt) p53 gene. The two R3 cell lines, DU145 (mt p53) and LNCaP (wt p53), are radiosensitive and show very different biochemical radiation response signatures from their p53 counterparts (figure 7.2, table 7.1).

Normally functioning wt p53 is required for G1 phase cell cycle arrest post-irradiation [49]. In this study, the measured fraction of G1 cells at 24 hours post-irradiation (table 7.2) confirms that the mt p53 cell lines do not show G1 phase arrest (6-11% G1 fraction at 24 hours post-irradiation) and in fact show high levels of radiation induced G2 phase arrest (60-87% G2 fraction at 24 hours post-irradiation). As expected, the wt p53 cell lines do show G1 phase arrest (40-73% G1 fraction at 24 hours post-irradiation).

7.4 Discussion: Study #1

7.4.1 RS detection of biochemical signatures of radiation response

The methods applied in this study, namely the acquisition of hundreds of high-quality single-cell RS spectra per cell line, established spectral processing techniques (chapters 4, 5 and 6), and PCA, enable the separation of radiation induced spectral changes

Table 7.2: RS biochemical radiation response category (RS Cat.), tissue of origin (TOI), percent variance explained by radiation induced PCA component, G1, S and G2 fractions at 24 hours post-irradiation, p53 status and average reported radiosensitivity (SF_2) for the seven cell lines used in this study.

RS Cat.	**Cell Line**	**TOI**	**Variance** (%)	**G1**, **S**, **G2** (% **at 24 hrs**)	**p53 Status**	**SF_2**
R1	H460	Lung	50.3	73, 10, 17	wt[196]	0.64[55, 183–185]
	MCF7	Breast	7.3	40, 9, 51	wt[197]	0.64[26, 55, 183, 186, 187]
	PacMet	Prostate	12.6	68, 7, 25	wt[150]	unknown
R2	MDA-MB-231	Breast	4.7	11, 10, 79	mt[197]	0.71[26, 55, 188, 189]
	PC3	Prostate	3.0	7, 6, 87	mt[198]	0.64[55, 190–192]
R3	DU145	Prostate	12.8	6, 34, 60	mt[198]	0.49[55, 190, 193–195]
	LNCaP	Prostate	4.0	64, 14, 22	wt[198]	0.27[190, 191, 194, 195]

from other simultaneously occurring sources of spectral variability, such as cell cycle. This technique provides direct analysis of the biomolecular changes arising in single cells responding to radiation exposure, independent of cell cycle or cell death related processes. RS signatures of radiation response (*i.e.*, the radiation induced PCA components, figure 7.2) are identified by statistically significant shifts in the corresponding PCA score distributions (*e.g.*, figures 7.3b & 7.4b, 7.3d & 7.4d and 7.3f & 7.4f) as a function of time post-irradiation and (in some cell lines) the delivered dose. The biomolecules responsible for these signatures are identified by the known molecular assignments of the positive and negative features in the radiation induced PCA components (figures 7.1a, 7.3a & 7.4a, 7.3c & 7.4c and 7.3e & 7.4e). In this study, it is found that the biomolecular radiation responses of the seven cell lines segregate into distinct categories (R1, R2 and R3), observable both by visual inspection of the signatures (figure 7.2) and by correlation analysis (table 7.1). The qualitative similarities observed in the biomolecular changes between categories R1 and R2, and the many opposite changes observed between category R3 and categories R1 and R2 (*i.e.*, figures 7.3a & 7.4a, 7.3c & 7.4c and 7.3e & 7.4e), are quantitatively confirmed via correlation analysis (table 7.1).

7.4.2 Segregation of common radiation response signatures according to p53 status and radiosensitivity

Of the five cell lines that fall into the RS radiation response categories R1 and R2, four are known to be radiation resistant ($SF_2 > 0.6$) and the PacMet radiosensitivity is unknown. The two cell lines in category R3 are comparatively radiation sensitive ($SF_2 < 0.5$) (table 7.2). The RS signatures of radiation response between categories R1 (resistant wt p53) and R2 (resistant mt p53) are different, but share many common molecular features (figure 7.2), indicating similar yet unique radiation responses. As both groups R1 and R2 are comprised solely of radiation resistant cell lines, it is possible that the biochemical radiation responses observed with RS are caused by cellular response mechanisms that increase survival after radiation exposure. Candidate mechanisms for such responses are discussed below (section 7.4.3).

The available SF_2 data from different laboratories (*e.g.*, for MCF7 the reported SF_2 values range from 0.50 [187] to 0.81 [183]) separate the confirmed resistant cell lines (average $SF_2 > 0.6$) from the comparatively sensitive cell lines (average $SF_2 < 0.5$). Reported SF_2 values were used if the clonogenic survival assay approximated

currently accepted protocols [27]. In particular, it was ensured that cultures were irradiated in exponential growth phase and that experiments were performed at least in triplicate.

It is important to note that the PCA calculation also provides the relative "strength" of the RS observed biochemical radiation response, given by the variance explained by the radiation induced PCA component identified. In this study, the cell lines demonstrated different radiation response strengths between categories of radiation response, and also within each category (figure 7.2, table 7.2). Between categories R1 and R2, the resistant wt p53 cell lines (R1) responded stronger than the resistant mt p53 cell lines (R2). Within each category, H460 responded stronger than PacMet which responded stronger than MCF7 (R1), MDA-MB-231 responded stronger than PC3 (R2), and DU145 responded stronger than LNCaP (R3). The relative strength of the biochemical radiation response may have some correlation to differences in radiosensitivity within a given RS category.

7.4.3 Biochemical mechanisms of radiation resistance or sensitivity

Radioresistant cell lines: RS categories R1 and R2

Some of the most dramatic radiation induced biochemical changes observed with RS in this study arise from proteins (figures 7.3 and 7.4). Between 24 and 72 hours post-irradiation, the known radiation resistant cell lines in RS categories R1 and R2 demonstrate increased concentrations of aromatic amino acids (phenylalanine, tyrosine and tryptophan) and α-helix protein structures, and decreased concentrations of β-sheet and random coil protein structures, relative to unirradiated controls. These changes corroborate with prior evidence that synthesis and degradation of structured proteins is correlated with increased survival post-irradiation and plays an important role in cellular radioadaptive responses [37]. These changes may also be reflective of cellular survival mechanisms triggered by radiation induced oxidative stress, involving the breakdown of structured proteins (*i.e.*, β-sheet and random coil protein structures) into free amino acids to aid in the scavenging of reactive oxygen species created by radiation damage [199]. The observation of increased α-helix protein structure with radiation could be explained by concurrent synthesis of certain proteins involved in radiation response pathways that increase cell survival post-irradiation. For example, colorectal cells have been shown to express 14-3-3σ protein for up to 60 hours post-

irradiation to aid in the inhibition of cell cycle progression through G2/M phase, resulting in increased survival [200]. 14-3-3 proteins are primarily composed of α-helices [201], are involved in many cell division and signalling pathways [201, 202], and are known to suppress apoptosis via inhibition of several pro-apoptosis pathways [202]. Another candidate protein with predominantly α-helix structure is survivin [203], a known anti-apoptosis factor that has been linked to increased radioresistance in glioblastomas [204], pancreatic cancers [205], rectal cancers [206] and head and neck squamous carcinomas [207]. Furthermore, it was shown for glioblastomas that radioresistant cell lines expressed survivin post-irradiation in all phases of the cell cycle, whereas radiosensitive cell lines limited survivin expression to G2/M phase, independent of p53 status [204].

In this study, similar radiation response signatures are observed between the resistant wt p53 (R1) and the resistant mt p53 (R2) cell lines. The p53 gene is known to regulate cell cycle arrest post-irradiation [49, 200] and flow cytometry measurements performed in this study (table 7.2) confirm the differences in cell cycle regulation that occur between the wt and mt p53 cell lines. Furthermore, radiation induced apoptosis, which would promote cell death post-irradiation and therefore increase sensitivity, requires wt p53 status [49]. As such, it is to be expected that the biochemical nature of an induced radiation response that promotes cell survival post-irradiation may be determined by p53 status. An example of such p53 dependence on biochemical radiation response and radiosensitivity has been recently observed via analysis of micro-RNA expression post-irradiation [208]. It is also likely that different responses working to achieve the same result (*i.e.*, increased survival post-irradiation) would have common characteristics (*e.g.*, expression of anti-apoptosis factors or other survival signals). As discussed above in section 7.4.2, this study observes stronger radiation responses from the R1 (resistant wt p53) cell lines than from the R2 (resistant mt p53) cell lines, although the responses have many similar features. If the observed RS responses from the R1 and R2 cell lines are indeed caused in part by the radiation induced synthesis of anti-apoptosis proteins, it is plausible that a wt p53 cell line with equivalent radiosensitivity as a mt p53 cell line would need to mount a larger biochemical response (*i.e.*, increased expression of anti-apoptosis factors) in order to suppress the intact pro-apoptotic pathways, in addition to the biochemical response required for other mechanisms of survival post-irradiation. This model may explain the observed differences and similarities between the radiation response signatures obtained for groups R1 and R2 in this study.

Radiosensitive cell lines: RS category R3

Neither of the radiosensitive cell lines comprising RS category R3, DU145 (mt p53) and LNCaP (wt p53), show a radiation response similar to the other five cell lines. This lack of response may be the result of some other reason why these cells do not mount a similar radiation response as their p53 counterparts, which may be necessary for increased survival post-irradiation (*e.g.*, expression of anti-apoptosis factors or other survival signals). Furthermore, both of the radiation response signatures for the R3 cell lines, although different (figure 7.2), are anti-correlated with the R1 and R2 signatures and are positively correlated with each other (table 7.1). As such, these R3 radiation response signatures, which show anti-correlation with possible biochemical signatures of radiation resistance, may in turn prove to be signatures of radiation sensitivity.

7.4.4 Uniqueness of the observed RS biochemical radiation responses

The biochemical variability described by the radiation induced PCA components (figure 7.2) is only observed when the data sets containing both unirradiated and irradiated cell spectra are input into PCA, with one exception. The radiation induced PCA component for the H460 cell line, which demonstrated the strongest radiation response of all the cell lines (50.3% of the total variance), is also observed ($r = 0.80$) when only the unirradiated H460 cell spectra are input into PCA but with a much reduced percent variance explained (16.8% vs. 50.3%). These observations suggest that the biochemical variability described by the radiation induced PCA component for the H460 cell line is not a uniquely radiation induced response, but rather is enhanced by radiation and may be induced by other factors as well. Interestingly, a large fraction of the variability described by the component obtained from the unirradiated cells data set arises from a small number of outliers ($\sim$4) from the 80 unirradiated cells, which can be seen in the unirradiated sample PCA score distributions for the H460 radiation induced PCA component shown in figure 7.3b. When these 4 outliers are removed from the 80 cell data set prior to PCA, the component is still observed but with a significant drop in variance (11.1% vs. 16.8%). For the other six human tumour cell lines investigated in this study, the radiation induced PCA components are not observed if only the unirradiated cell spectra are input into PCA. If the biochemical variability is indeed present in these other cell lines it is too weak for

detection with RS using the current methods, possibly due to insufficient induction by the *in vitro* culture environment. Since the *in vitro* radiation response of the H460 cell line is much stronger than that of the other cell lines, it is possible that the *in vitro* culture environment provided sufficient stimuli or stresses to induce a similar biochemical response in a subset of the unirradiated H460 cells, which was therefore detectable with RS without radiation exposure.

7.4.5 Effect of radiation on cell cycle variability

In chapter 6 it is shown that for RS of irradiated DU145 cells, radiation induces a observable effect on the PCA scores for the cell cycle PCA component, while leaving the features of the cell cycle PCA component unchanged. Here, the same result is obtained for all seven cell lines used in this study (figure 7.1d and table 7.2). The effect of radiation on the corresponding PCA scores for the cell cycle PCA component (not shown) is dependent on the susceptibility of the cell line to radiation induced depletion of nucleic acid and protein content relative to lipid content, arising from early G1 (or G0) arrest or like processes (chapter 5 and corresponding publication [168]). Determining any possible relationships between this susceptibility and the nature of the observed RS signatures of radiation response, or with known radiosensitivity, would require similar analysis of more cell lines and may be a topic of interest for future research.

7.5 Materials & Methods: Study #2

7.5.1 TRAMP-C2 cells

TRAMP-C2 cells, originally obtained from American Type Culture Collection (ATCC), were cultured *in vitro*, using previously described protocols (section 3.1.2), with DMEM growth media (HyClone) supplemented with 5% FBS (HyClone), 5% Nuserum IV (Collaborative Biomedical Products, Bedford, MA) and 5 μg/mL insulin (HyClone).

The TRAMP-C2 cell line contains wt p53 [209]. Only one previous study has investigated the intrinsic radiosensitivity of the TRAMP-C2 cell line [210], from which a SF_2 value of $\sim$0.5 was estimated from the reported survival curve for comparisons with the other cell lines investigated here. This same study also tested the radiosensitivity of the LNCaP cell line and reported that TRAMP-C2 cells are approximately

twice as resistant as LNCaP cells [210]. The SF_2 value of ∼0.5 for the TRAMP-C2 cell line is therefore similar to twice the average SF_2 reported here for the LNCaP cell line ($SF_2 = 0.27$). As such, compared to the other wt p53 cell lines used in study #1 (classified into RS categories R1 or R3), the TRAMP-C2 cell line is more resistant than the LNCaP cell line (R3, $SF_2 = 0.27$) but is likely more sensitive than the H460 and MCF7 cell lines (R1, $SF_2 = 0.64$).

7.5.2 *In vitro* experiment

Irradiation of *in vitro* cell cultures, cell cycle and viability analysis, and RS acquisition was performed as described above for the human tumour cell lines (study #1). For the *in vitro* TRAMP-C2 experiment presented below (section 7.6.1), the viability of all cultures was $\geq 94\%$.

7.5.3 *In vivo* experiment

All mouse handling, irradiation, and solid tumour harvesting was performed by trained members of the RCIT group following protocols approved by the University of Victoria's Animal Care Committee. Approximately 3×10^6 TRAMP-C2 cells were injected into the subcutaneous layer of the flanks of two mice and tumour growth was monitored until tumours grew to ∼10 mm in diameter (∼9 weeks). Mice were irradiated under dissociative anaesthesia, induced by a 200/0.1 mg/kg mixture of Ketamine/Domitor in 400 μL sterile water injected subcutaneously. No effect on tumour oxygenation or radiation response is expected as a result of the anaesthetic. The tumour of one mouse was irradiated to 27.2 Gy over 5 days according to the following protocol: 6 Gy on days 1 and 2, no dose on day 3, and 7.6 Gy on days 4 and 5. The second mouse tumour was sham-irradiated to serve as a control. Mice were euthanized ∼24 hours after delivery of the final fraction of radiation and solid tumours were immediately removed and stored on ice in growth media until processing.

Tumour cell suspensions were extracted from the solid tumours in a series of processing steps. The solid tumours were first cut and mashed in a 10 cm dish containing ∼2 mL of Collagenase D (Roche Diagnostics, Laval, QC, Canada) for 5 minutes, then transferred to a 15 mL tube with ∼5 mL Collagenase and incubated in a 37 °C water bath for 30 minutes. The mixture was then pipetted vigorously to further disassociate tumour cells from other tissues, passed through a nylon screen to separate out any skin tissues and fatty deposits, and then washed twice with PBS. To isolate

tumour cells from other tumour infiltrating cells (*i.e.*, red and white blood cells), the suspension was placed on an 80%/40% gradient of Percoll-Plus (GE Healthcare Bio-Sciences, Piscataway, NJ, USA) diluted in growth media and centrifuged for 15 min at 1800 rotations per minute to form separate layers of tumour cells and tumour infiltrates. The tumour cell layer was extracted with a pipet tip and the suspension was then washed twice in PBS, resuspended in growth media, centrifuged into a 200 μL vial, and stored on ice until RS analysis.

The cell cycle distribution and viability of the unirradiated and irradiated samples was determined with flow cytometry; the viability of the unirradiated and irradiated samples was found to be 61% and 71%, respectively. RS acquisition was performed as previously described, except that only LWN window Raman spectra were acquired. Spectra were obtained from 40 cells for each sample, resulting in a total of 80 spectra for the *in vivo* TRAMP-C2 data set.

7.6 Results & Discussion: Study #2

7.6.1 *In vitro* experiment

Unirradiated TRAMP-C2 cell spectra

The averaged Raman spectrum for 20 unirradiated cells from the TRAMP-C2 cell line, harvested and analyzed immediately after the time of irradiation, yields the same features, at comparable intensities, as the averaged spectra for the previously studied human tumour cell lines (figure 7.5). As before (figure 7.1b), there are subtle differences in the relative intensity of spectral features between the TRAMP-C2 averaged spectrum and the averaged spectra for the other cell lines, arising from inherent differences in biomolecular compositions between tumour cell lines [79] and slightly different cell cycle distributions between cultures at time of irradiation [168].

PCA results

The first three PCA components and the corresponding PCA score plots for the 200 cells in the *in vitro* TRAMP-C2 irradiation study are presented in figure 7.6, with the percent variance explained by each component provided. The remaining PCA components each explain less than 2% of the total variance and do not display any measurable trend in the score plots, suggesting that the meaningful biochemical

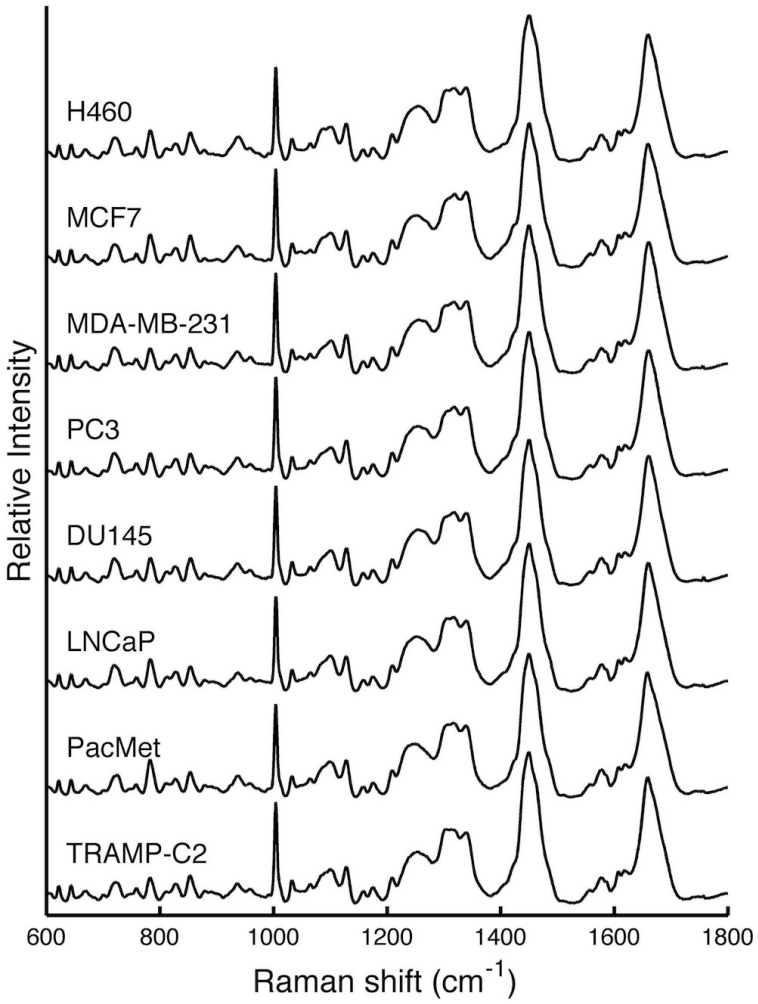


Figure 7.5: Averaged spectra from 20 unirradiated cells for the seven human tumour cell lines used in study #1 and the TRAMP-C2 mouse prostate tumour cell line used in study #2.

variability detectable with RS for the *in vitro* TRAMP-C2 irradiation experiment is contained in the first three PCA components.

The first PCA component (figure 7.6a) is the same cell cycle component as previously observed for the human tumour cell lines (figures 7.1b & 7.1d) and explains a comparable amount of the total variance in the data set (30.7%). The PCA scores for the first PCA component (figure 7.6b) demonstrate a similar trend to lower PCA scores (*i.e.*, decreased protein and nucleic acid content relative to lipid content) with increased time after sub-culturing as is observed for DU145 cells (figures 5.4a and 6.2b) and for the other human tumour cell lines (data not shown).

The second PCA component (figure 7.6c) explains 3.9% of the total variance and is dominated by a feature at $\sim$1660 cm^{-1} corresponding to protein amide groups. The PCA scores do not show any consistent separations between unirradiated and irradiated samples. This PCA component is also obtained during analysis of the TRAMP-C2 data set comprised of only unirradiated samples; analysis of the PCA scores from that data set (not shown) indicates that this component arises from variability correlated with culture confluency (or time since sub-culturing), similar to what is presented in chapter 5 for the DU145 cell line.

The third PCA component (figure 7.6e) explains 2.4% of the total variance. This component shares many features with the radiation induced PCA components obtained from the R1 and R2 human tumour cell lines. The common negative features arise from amino acids and α-helix protein structures and the common positive features arise from nucleic acids, CH_2 groups and β-sheet protein structures. The PCA scores for the third PCA component (figure 7.6f) show a slight yet consistent decrease in the average PCA score values for irradiated cultures as compared to unirradiated controls from 24 to 72 hours post-irradiation, although only the decrease for the 72 hours 30 Gy culture is statistically significant ($p = 0.01$). However, this third PCA component is also obtained during analysis of the *in vitro* TRAMP-C2 data sets comprised of only unirradiated samples ($r = 0.81$) and only irradiated samples ($r = 0.92$). This indicates that the biochemical variability described by this PCA component is not uniquely induced by radiation exposure, a result that is also obtained for the H460 cell line (section 7.4.4). When the TRAMP-C2 data set containing only unirradiated samples is input into PCA, there are no significant changes in the PCA score distributions between unirradiated samples at any time points. When the TRAMP-C2 data set containing only irradiated samples is input into PCA, there are significant decreases ($p < 0.01$) in the PCA score distributions for the irradiated samples at 24

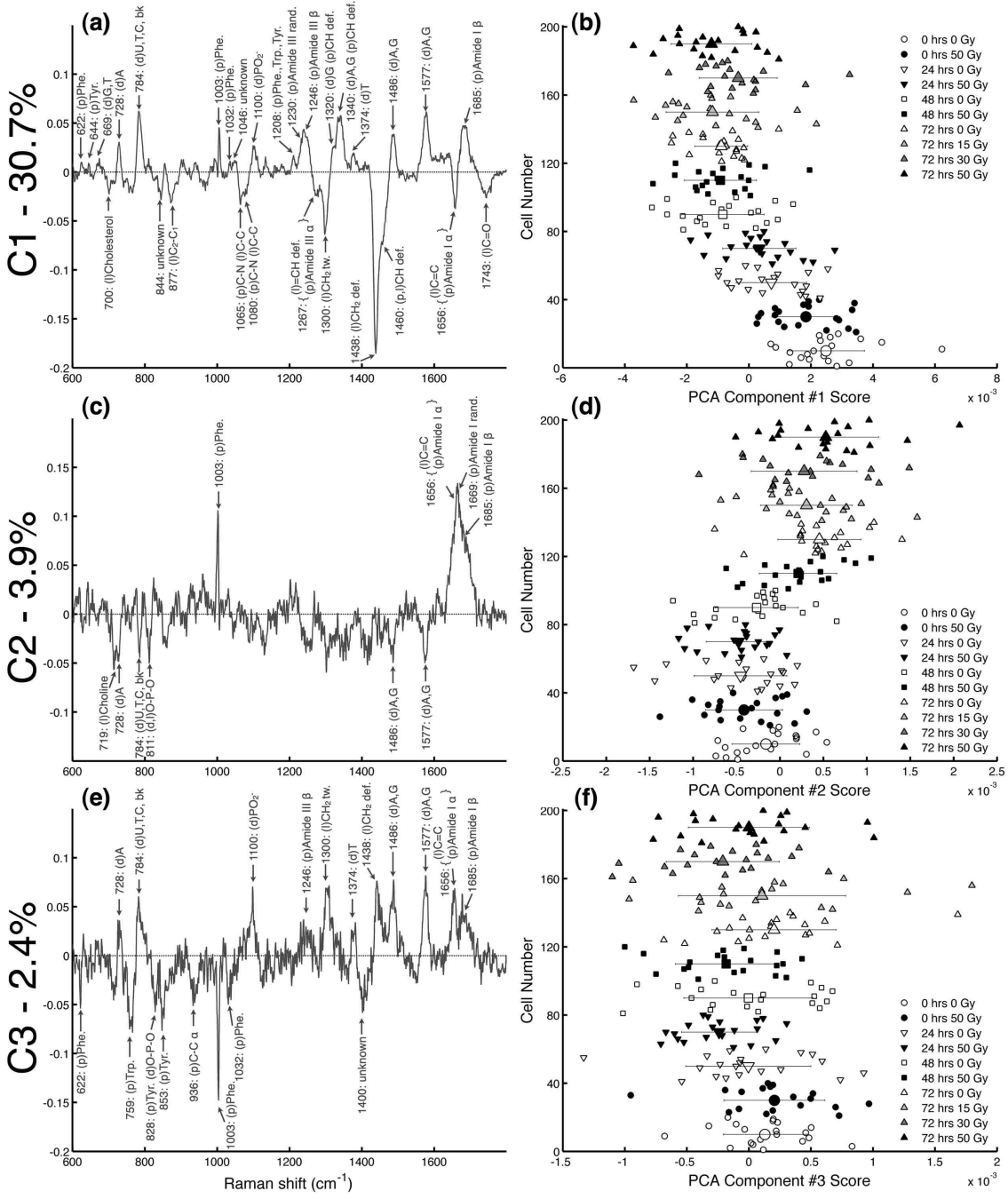


Figure 7.6: (a, c, e) First three PCA components (C1, C2, C3) and (b, d, f) PCA score plots for the *in vitro* TRAMP-C2 irradiation experiment. (a, c, e) Raman shifts and molecular origins of identifiable features, and the percent variance explained, are provided. Abbreviations - as in figure 7.1. (b, d, f) Different markers categorize all 200 cells by time of RS acquisition after irradiation. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

and 48 hours post-irradiation with respect to the 0 hours post-irradiation sample. As a result of this analysis and due to the similarities with the previously observed radiation induced PCA components (corroborated below by correlation analysis), the third PCA component observed here for the TRAMP-C2 cell line is tentatively assigned as a radiation induced PCA component for the following analyses.

The cell cycle and radiation induced PCA components obtained from the TRAMP-C2 irradiation experiment (figures 7.6a and 7.6e) are compared with the corresponding components obtained from the human tumour cell lines in figures 7.7a and 7.7b, respectively. Correlation analysis between the TRAMP-C2 radiation induced PCA component and the radiation induced PCA components obtained from the human tumour cell lines indicates that the TRAMP-C2 radiation induced PCA component is indeed similar to the radiation response signatures obtained from the R1 (H460, MCF7 and PacMet) and R2 (MDA-MB-231 and PC3) cell lines in study #1 ($0.29 < r < 0.64$). Furthermore, there is significant anti-correlation between the TRAMP-C2 component and the DU145 component ($r = -0.33$) and no significant correlation with the LNCaP component ($r = 0.01$).

TRAMP-C2 *in vitro* radiation response category

The TRAMP-C2 *in vitro* radiation response observed with RS is very weak, as indicated by the low percent variance explained by the radiation induced PCA component (2.4%) and the limited significance of the changes in the corresponding PCA score distributions (figure 7.6f). As such, the TRAMP-C2 radiation induced PCA component is quite noisy compared to the components obtained previously (figure 7.7b), therefore suppressing the r-values obtained during correlation analysis and introducing uncertainty into the assignment of the TRAMP-C2 radiation response into one of the currently defined radiation response categories: R1, R2 or R3 (although the TRAMP-C2 response is clearly quite different from the responses of the R3 category cell lines). However, among categories R1 and R2 the highest r-value ($r = 0.64$) is obtained with the PacMet cell line (category R1) and the lowest r-value ($r = 0.29$) is obtained with the MDA-MB-231 cell line (category R2). This suggests that the radiation response of the TRAMP-C2 line is most similar to the radiation response of the R1 cell lines, while also similar to the radiation responses of the R2 cell lines (a result consistent with the positive correlations between categories R1 and R2 in study #1). However, the strength of the TRAMP-C2 radiation response (2.4%) is much weaker than the responses of the other R1 cell lines (7.3%, 12.6% and 50.3%)

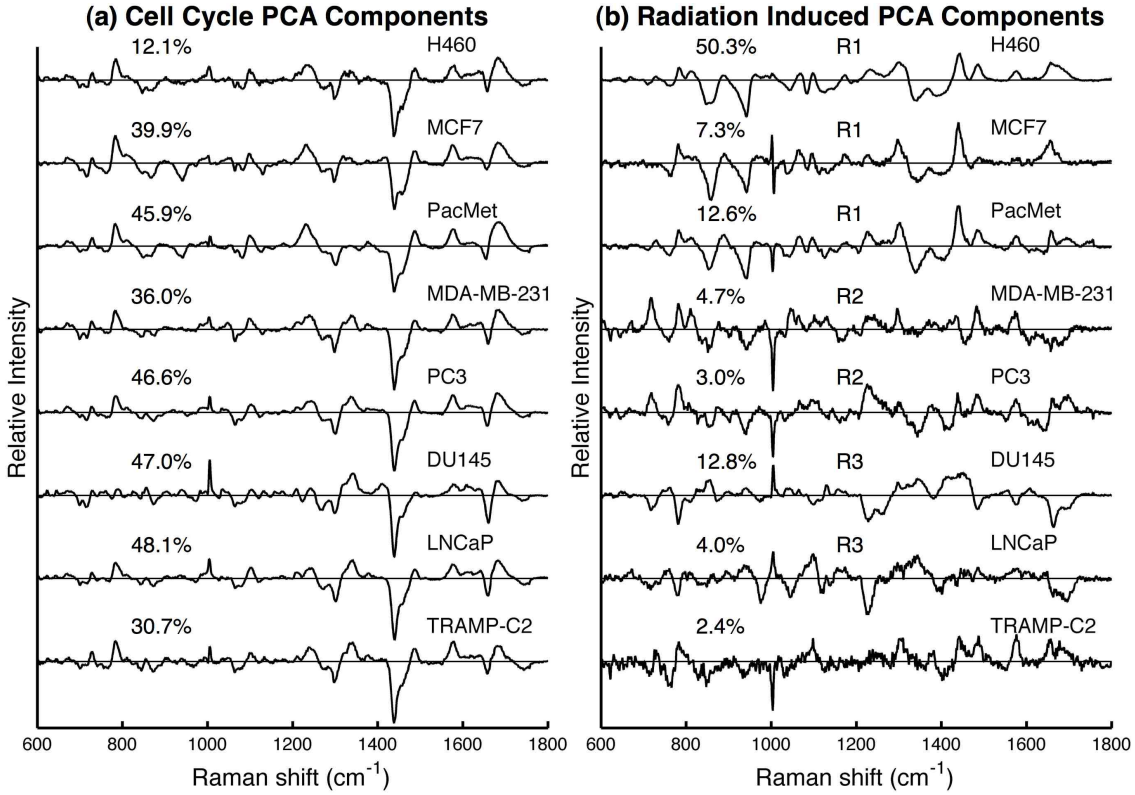


Figure 7.7: (a) Cell cycle PCA components and (b) radiation induced PCA components for the TRAMP-C2 mouse prostate tumour cell line and the seven human tumour cell lines, with the previously assigned RS radiation response categories indicated in (b). For TRAMP-C2, the Raman shift and molecular origin of identifiable features are provided in figures (a) 7.6a and (b) 7.6e.

and is slightly weaker than the responses of the R2 cell lines (4.7% and 3.0%) (figure 7.7b).

Biochemical interpretation of TRAMP-C2 *in vitro* radiation response

The relationship between the RS radiation response categories (R1, R2, R3) and tissue of origin, percent variance explained by the radiation induced PCA component, G1, S and G2 fractions at 24 hours post-irradiation, p53 status and average reported radiosensitivity is shown in table 7.2 (section 7.3.5) for the human tumour cell lines investigated in study #1. The data is reproduced below (table 7.3), with the *in vitro* TRAMP-C2 data included in italics for comparison. The TRAMP-C2 cell cycle distribution at 24 hours post-irradiation is consistent with the distributions for the other human tumour cell lines containing wt p53. The accuracy of the single SF_2 value reported for the TRAMP-C2 cell line is unknown, although the value ($SF_2 = 0.5$) suggests that the TRAMP-C2 cell line is neither overly radioresistant nor overly radiosensitive. The SF_2 value is comparable to that of the DU145 cell line ($SF_2 = 0.49$), although the DU145 cell line contains mt p53 and therefore is expected to respond differently, even with an approximately equivalent SF_2 value. If the radiation response signatures observed with RS for the R1 and R2 category cell lines are indeed the result of biochemical response mechanisms that increase survival post-irradiation (as postulated in section 7.4.3), it is possible that the reduced strength of the *in vitro* TRAMP-C2 radiation response observed with RS is the result of the reduced radioresistance of the TRAMP-C2 cell line relative to the other radioresistant wt p53 cell lines assigned to category R1 (table 7.3). The lack of an *in vitro* TRAMP-C2 radiation response that correlates with the response of the LNCaP cell line (wt p53, $SF_2 = 0.27$, RS category R3) may in turn indicate a lack of biochemical mechanisms associated with further reductions in radiation survival.

7.6.2 *In vivo* experiment

Averaged Raman spectra

The averaged Raman spectra for 20 unirradiated and irradiated cells from the TRAMP-C2 tumours grown and irradiated *in vivo* are shown in figure 7.8, along with a pair of averaged Raman spectra from the *in vitro* TRAMP-C2 experiment indicating the range of differences observable in the averaged spectra between *in vitro* samples ('0 hrs 0 Gy' vs. '72 hrs 0 Gy'). The spectra obtained from the *in vivo* tumour cells contain

Table 7.3: Comparison of the TRAMP-C2 *in vitro* RS biochemical radiation response category (RS Cat.), tissue of origin (TOI), percent variance explained by radiation induced PCA component, G1, S and G2 fractions at 24 hours post-irradiation, p53 status and average reported radiosensitivity (SF_2) with the seven human tumour cell lines used in study #1. TRAMP-C2 values are italicized for clarity.

RS Cat.	**Cell Line**	**TOI**	**Variance** (%)	**G1**, **S**, **G2** (% **at 24 hrs**)	**p53 Status**	**SF_2**
R1	H460	Lung	50.3	73, 10, 17	wt[196]	0.64[55, 183–185]
	MCF7	Breast	7.3	40, 9, 51	wt[197]	0.64[26, 55, 183, 186, 187]
	PacMet	Prostate	12.6	68, 7, 25	wt[150]	unknown
	TRAMP-C2	*Prostate*	*2.4*	*58, 15, 27*	*wt*[209]	*∼0.5*[210]
R2	MDA-MB-231	Breast	4.7	11, 10, 79	mt[197]	0.71[26, 55, 188, 189]
	PC3	Prostate	3.0	7, 6, 87	mt[198]	0.64[55, 190–192]
R3	DU145	Prostate	12.8	6, 34, 60	mt[198]	0.49[55, 190, 193–195]
	LNCaP	Prostate	4.0	64, 14, 22	wt[198]	0.27[190, 191, 194, 195]

the same features, at comparable intensities, as the averaged spectra obtained from the *in vitro* experiment. There are several wavenumbers where consistent differences in the relative intensities of spectral features can be observed between the *in vitro* and *in vivo* averaged spectra, which are further analyzed in section 7.6.3 below.

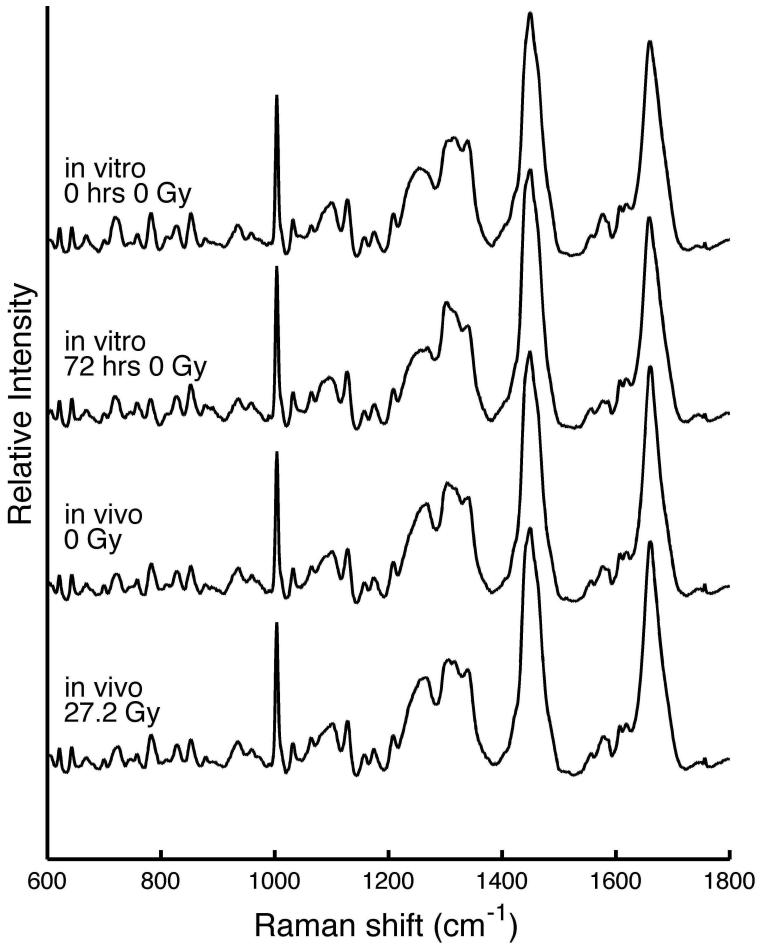


Figure 7.8: Averaged spectra from 20 unirradiated and 20 irradiated cells from the TRAMP-C2 tumours grown and irradiated *in vivo*, with a pair of averaged Raman spectra from the *in vitro* TRAMP-C2 experiment indicating the range of differences observable in the averaged spectra between *in vitro* samples ('0 hrs 0 Gy' vs. '72 hrs 0 Gy').

PCA results

The first two PCA components from the *in vivo* TRAMP-C2 irradiation experiment explain 43.0% and 8.0% of the total variance. The third PCA component explains 4.6% of the total variance and arises from variability in the quartz substrate. The re-

mainder of the components each explain less than 2.5% of the total variance. As such, the biochemical variability detectable with RS in the *in vivo* TRAMP-C2 irradiation experiment is likely contained in the first two PCA components.

The first PCA component is shown in figure 7.9a, with the first PCA component from the *in vitro* TRAMP-C2 irradiation experiment included for comparison. The first PCA component from the *in vivo* data set is the same cell cycle component as is previously observed for all cell lines cultured *in vitro*, with a high positive correlation ($r = 0.92$) with the cell cycle component obtained from the *in vitro* TRAMP-C2 experiment (figure 7.9a). The *in vivo* PCA scores for the first PCA component (figure 7.9b) indicate a significant ($p = 0.001$) increase in PCA score values for the irradiated tumour cells, reflecting increased levels of protein and nucleic acid relative to lipid content that likely indicates an increase in tumour cell proliferation relative to the unirradiated tumour cells. This result is supported by the cell cycle distributions obtained with flow cytometry for these samples, which indicate an increase in the combined S and G2 fraction from 33% for the unirradiated tumour cells to 45% for the irradiated tumour cells. This increased proliferation, detected with both RS and flow cytometry, may be the result of tumour repopulation in response to the fractionated radiation treatment [31], or may indicate inherent differences in the tumour microenvironment between the tumours of the two mice. The latter explanation is highly probable, as it was found that the tumour selected for irradiation was larger, more vascularized, and contained a higher number of tumour infiltrating white blood cells (personal correspondence with Dr. Lisa Johnson, post-doctoral fellow with BCCA-VIC). These factors suggest that the irradiated tumour cells had increased access to nutrients provided by the tumour vasculature, allowing for increased growth and proliferation relative to the unirradiated tumour cells.

The second PCA component from the *in vivo* TRAMP-C2 irradiation experiment is shown in figure 7.9c. This component contains many similar features as the third PCA component from the *in vitro* TRAMP-C2 irradiation experiment (also shown in figure 7.9c for comparison), which is tentatively assigned as a radiation induced PCA component. Furthermore, there is a high positive correlation ($r = 0.58$) between the third *in vitro* PCA component and the second *in vivo* PCA component. However, for the *in vivo* TRAMP-C2 experiment the second PCA component is also observed when only the unirradiated tumour cells are input into the PCA computation, but with a reduced percent variance explained (6.5% for only unirradiated cells versus 12.4% for only irradiated cells). As such, the biochemical variability described by this PCA

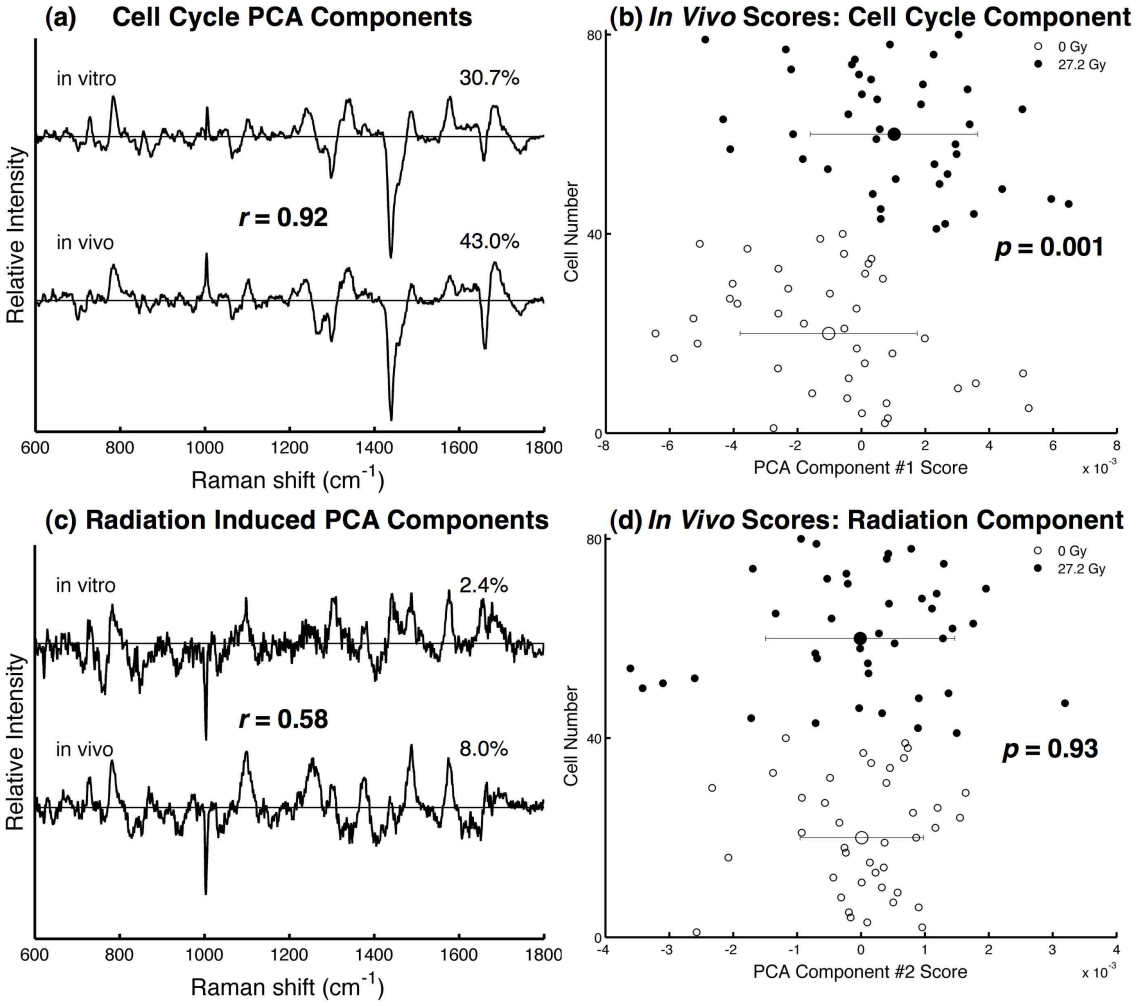


Figure 7.9: (a) Cell cycle and (c) radiation induced PCA components, with the percent variance explained, and (b, d) the corresponding PCA scores, for the *in vivo* TRAMP-C2 irradiation experiment. The corresponding PCA components from the *in vitro* TRAMP-C2 experiment are included in (a) and (c) for comparison. (a, c) Correlations between components (r-values) and (b, d) statistical significances of differences between *in vivo* PCA score distributions (p-values) are provided.

component is not unique to the irradiated tumour cells, although the amount of variability may be enhanced by irradiation. The PCA scores for the second *in vivo* PCA component (figure 7.9d) corroborate this result. There is no significant shift in the PCA score distributions between the irradiated and unirradiated tumour cells ($p = 0.93$), but there is a $\sim$50% increase in the standard deviation of the irradiated PCA scores relative to the unirradiated PCA scores (0.00148 vs. 0.00097).

Uniqueness of the TRAMP-C2 radiation response

The biochemical variability described by the second *in vivo* PCA component (figure 7.9b) is also observed when only the unirradiated cell spectra are input into PCA, as is also observed *in vitro* (section 7.6.1), and is therefore not a unique radiation response. However, it is evident that the biochemical variability may be enhanced by radiation exposure. This result is in contrast with most of the *in vitro* irradiation experiments previously presented (study #1), where the radiation induced PCA components are not observed if only the unirradiated data set is input into PCA. However, this result is also observed for the H460 cell line (recall section 7.4.4), where the radiation induced PCA component is also detected if only the unirradiated data set is input into PCA but with a large decrease in percent variance explained (16.8% vs. 50.3%). This suggests that the biochemical variability described by the radiation induced PCA components for both the H460 and the TRAMP-C2 cell lines are not uniquely radiation induced responses, but are rather enhanced by radiation and may be induced by other factors as well. If, as postulated in section 7.4.3, the category R1 and R2 radiation responses observed with RS are indeed the result of increased expression of anti-apoptosis factors or other survival signals post-irradiation, it is possible that the *in vivo* tumour microenvironment may cause the induction of a biochemical stress response similar to what is observed post-irradiation. This theory may explain the increase in percent variance explained by the *in vivo* TRAMP-C2 radiation induced component, as compared to the corresponding component obtained *in vitro* (8.0% vs. 2.4%, figure 7.9b).

7.6.3 Combined *in vitro* and *in vivo* data set analysis

Concatenating the *in vitro* and *in vivo* TRAMP-C2 irradiation experimental data sets into a single 280 spectra data set prior to PCA calculation allows for direct comparisons of the sources of biochemical variability between the *in vitro* and *in vivo* cells.

The first three PCA components and the corresponding PCA score plots for the combined *in vitro* and *in vivo* TRAMP-C2 irradiation experiments data set are presented in figure 7.10, with the percent variance explained by each component provided. The fourth PCA component from this data set explains 2.2% of the total variance and arises from variability in quartz substrate signals. The remaining PCA components each explain less that 2% of the total variance, suggesting that the biochemical variability detected with RS in the combined data set is contained in the first three PCA components.

The first PCA component from the combined data set (figure 7.10a) is the same cell cycle PCA component as previously observed and explains 28.1% of the total variance. The PCA scores for the first PCA component (figure 7.10b) demonstrate the same trend to lower PCA scores (*i.e.*, decreased protein and nucleic acid content relative to lipid content) with increased time after sub-culturing for the *in vitro* samples, as is observed for the *in vitro* data set (figure 7.6b). The PCA scores for the *in vivo* samples demonstrate the same significant ($p = 0.001$) shift towards higher PCA scores for the irradiated tumour cells relative to the unirradiated tumour cells, as is observed for the *in vivo* data set (figure 7.9b). Therefore, the biochemical interpretation and the discriminatory capability of the PCA results for analysis of cell cycle spectral variability is maintained when *in vitro* and *in vivo* data sets are combined in such fashion. It is worth noting that the range of the PCA score values for each of the *in vivo* samples encompasses the total range of the *in vitro* samples for all time points. This indicates that there is a greater range of biochemical variability related to cell cycle between tumour cells obtained *in vivo* than between tumour cells from a single culture *in vitro*. This is likely a result of the heterogeneity of the tumour microenvironment, specifically regarding access to nutrients required for proliferation.

The second PCA component from the combined data set (figure 7.10c) explains 21.6% of the total variance and is unique to the combined *in vitro* and *in vivo* data set. Many features can be assigned to known molecules, primarily from proteins, with a number of positive amino acid features of moderate intensity at 759, 853, 1003 and 1032 cm^{-1} and a pair of highly intense negative features at 1267 and 1663 cm^{-1} arising from amide groups in protein secondary structure. The PCA scores for the second PCA component (figure 7.10d) indicate that the spectral variability described by the second PCA component arises solely from spectral differences between the *in vivo* and *in vitro* samples, with a highly significant ($p < 0.0001$) shift to more negative PCA scores for the *in vivo* PCA score distributions. These differences can be directly

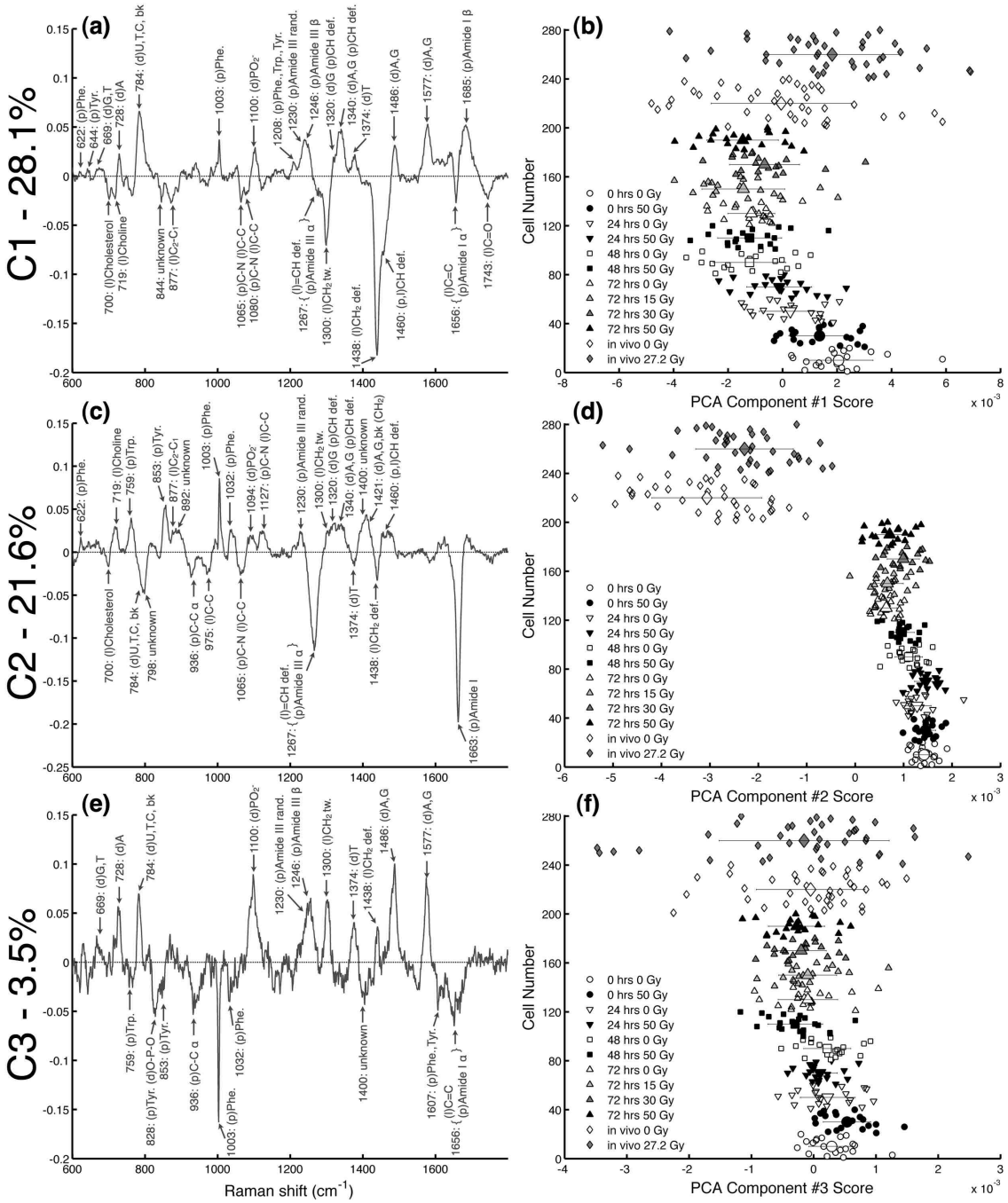


Figure 7.10: First three (a, c, e) PCA components (C1, C2, C3) and (b, d, f) PCA score plots for the combined *in vitro* and *in vivo* TRAMP-C2 irradiation experiments data set. (a, c, e) Raman shifts and molecular origins of identifiable features, and the percent variance explained, are provided. Abbreviations - as in figure 7.1. (b, d, f) Different markers categorize all 200 *in vitro* cells by time of RS acquisition after irradiation, with the 80 *in vivo* cells indicated by diamond (◇) markers. The average score and standard deviation is shown for each sample for visualization of the trends in the data.

observed in the averaged Raman spectra obtained from selected samples (recall figure 7.8), most notably at 1267 and 1663 cm^{-1} where the *in vivo* averaged spectra indicate increased intensities at these wavenumbers relative to the *in vitro* averaged spectra. The biochemical reason for these differences is currently unknown. The differences may arise due to the nature of the *in vivo* tumour microenvironment as compared to *in vitro* culture conditions, potentially from non-cellular sources contributing to the Raman spectra obtained from the cells grown *in vivo*. One possibility is the presence of residual collagen content from the *in vivo* tumour tissues, as the Raman spectrum of collagen (a protein constituent of extracellular connective tissue) contains several features arising from secondary protein structures [211]. However, not all features of a collagen spectrum [211] are reproduced in the second PCA component obtained here (figure 7.10c). Alternatively, the differences may arise from the additional processing techniques used for extraction of the tumour cell suspension from the solid tumours (recall section 7.5.3). The spectral variability described by this second PCA component is not a significant source of spectral variability between the *in vivo* tumour cells alone, as this component is not observed during independent analysis of the *in vivo* TRAMP-C2 irradiation experiment (figure 7.9).

The third PCA component from the combined data set (figure 7.10e) explains 3.5% of the total variance. This component is almost identical to the PCA component tentatively assigned as a radiation induced PCA component during independent analysis of the *in vivo* TRAMP-C2 experiment (figure 7.9c, bottom trace) ($r = 0.89$) and is similar to the corresponding weak PCA component obtained from the *in vitro* TRAMP-C2 experiment (figure 7.9c, top trace) ($r = 0.59$). As observed previously for the *in vivo* experiment, the standard deviation of the irradiated PCA score distribution (figure 7.10f) is increased by $\sim$50% relative to the unirradiated distribution. Furthermore, as observed previously for the *in vitro* experiment, there is limited significance of the changes in the *in vitro* irradiated PCA score distributions with respect to unirradiated controls (figure 7.10f), with only the 48 hours 50 Gy sample returning a statistically significant ($p = 0.0003$) shift to lower PCA score values. Therefore, as concluded above for analysis of the first PCA component, the biochemical interpretation and the discriminatory capability of the PCA results for analysis of spectral variability is maintained when *in vitro* and *in vivo* data sets are combined in such fashion. In addition, the range of the PCA score values for each of the *in vivo* samples is larger than the total range of the *in vitro* samples (figure 7.10f), which is likely a further indication of the heterogeneity of the tumour microenvironment.

7.6.4 *In vivo* experimental difficulties and possible resolutions

Current difficulties

The two main difficulties that arose during the *in vivo* experiment presented here were (1) low tumour cell yield and (2) low tumour cell viability. The currently implemented sample preparation techniques for single-cell RS acquisition require a sufficiently sized cell pellet to facilitate physical transfer of the cells to the quartz substrate for single-cell RS acquisition with a dry microscope objective. In addition, a cell pellet several cell layers thick is desirable to minimize background contributions from the quartz substrate. For the *in vivo* tumour cell samples collected in this work, the tumour cell yield was much reduced from that of the *in vitro* cultures, resulting in difficulties in transferring the cell pellets to the quartz substrate and therefore yielding small cell pellets on the substrate of only 1-2 cell layers thick. This introduced more spectral variability arising from the quartz substrate, but did not appear to negatively affect the analysis of the *in vivo* data set (figure 7.9) as the variability from the quartz substrate was restricted to a single PCA component explaining a low amount (4.6%) of the total variance. The low tumour cell viability (61% for unirradiated, 71% for irradiated) introduced difficulties during RS acquisition, as there was a corresponding increased incidence of the acquisition of a cell spectrum resembling that of a dead cell, thus requiring rejection of the spectrum and the acquisition of a new cell. Although not problematic from an analysis perspective, this process significantly lengthened the total time required for acquisition of cell spectra from the sample, possibly resulting in spectral variability arising from the extended exposure of the live cells to a non-aqueous environment. However, no such variability is observed in the *in vivo* experiment presented here.

Although both these difficulties were surmounted in this pilot study by careful sample preparations and a high level of diligence during RS acquisition, these difficulties may limit the robustness and portability of such RS techniques for future studies, if unresolved.

Possible resolutions

The *in vivo* tumour cell yield is fundamentally limited by the size of the tumour, but is also currently restricted by the tumour processing methods. The processing steps used in this work (section 7.5.3) induce significant tumour cell loss during disaggregation and filtration, but perhaps may be optimized for future work. However, the difficulties

imposed for RS acquisition by low tumour cell yield could likely be overcome by using a high power oil immersion microscope objective, with cells kept in an aqueous environment under a quartz cover slip. A typical 100x oil immersion objective has a numerical aperture (NA) of ∼1.3, which would provide a ∼50% improvement in confocal resolution relative to the 100x dry objective with a NA of 0.9 used in this work (recall the depth-of-focus formula, equation 2.17) and would possibly negate the current benefits of a thicker cell pellet. Furthermore, sample preparation in an aqueous environment would allow the use of a pipet to transfer small numbers of cells to the substrate. Currently the tip of a flat scalpel blade is used, which is limited to the transfer of physically manipulatable pellets.

The tumour cell viability is fundamentally limited by the original viability of the extracted tumour. However, biochemical techniques for separating live and dead cells exist that do not require cell fixation or staining; therefore, such techniques would not affect the Raman spectra of the cells. One recently developed technique is called contactless dielectrophoresis [212], which utilizes changes in the electrical conductivity of a dead cell, resulting from the loss of membrane integrity that occurs with cell death, to separate live cells from dead cells using electric field gradients. Such techniques would ensure the collection of only live cells for RS analysis, thus removing the current requirement for rejecting dead cell spectra during RS acquisition.

7.7 Conclusion

The primary results of this chapter (study #1) demonstrate that RS can detect biochemical signatures of *in vitro* radiation response that segregate according to p53 status and intrinsic radiosensitivity (SF_2), within a preliminary panel of seven human tumour cell lines derived from prostate (DU145, PC3, LNCaP and PacMet), breast (MDA-MB-231 and MCF7) and lung (H460). The observed RS signatures arise from radiation induced changes in cellular concentrations of aromatic amino acids, conformational protein structures, and certain nucleic acid and lipid functional groups, and are detected from live, unfixed, single cells analyzed 1-3 days post-irradiation. This sensitivity to the biomolecules responsible for the observed radiation responses provides new insight into possible mechanisms of radiation survival, and into the differences in such survival mechanisms between wt p53 and mt p53 cell lines. Preliminary results (study #2) obtained from a mouse prostate tumour cell line (TRAMP-C2), irradiated both *in vitro* and *in vivo*, indicate that the candidate RS signatures of ra-

diation resistance or sensitivity observed *in vitro* may also be detectable from tumour cells obtained from an *in vivo* system during radiation therapy treatment.

Chapter 8

Conclusions and Future Work

As stated in section 1.4, the primary goals of this work are (1) to develop a robust methodology for RS acquisition and data processing techniques for highly sensitive and specific spectral analysis of biochemical changes in single human cells, (2) to demonstrate the utility of such RS techniques for analysis of radiation induced biochemical changes in a human tumour cell line, and (3) to apply RS to investigate the biochemical radiation response of a panel of human tumour cell lines with varying intrinsic radiosensitivity. The achievement of these goals is demonstrated in this work by the results presented and discussed in chapters 4 to 7, which are summarized in turn below. The thesis then concludes with a discussion of some important implications of the results of this work for future studies.

Summary of key results

Chapter 4 presents an analysis of the data processing techniques implemented in this work, in order to demonstrate the robustness and validity of the developed techniques for applying PCA to large Raman data sets obtained from single cells. A representative set of raw data is used to demonstrate the effect of varying certain processing parameters (related to spectral smoothing, baseline estimation and spectral normalization) on the final outcomes and interpretations of the analysis. It is paramount in this work to have confidence in the outcomes of the data analysis, most importantly in (1) the detection of molecular sources of spectral variability as represented by the features in the PCA components, and (2) the discriminatory capability of the PCA results as represented by the separations between PCA score distributions for a given PCA component. The data processing steps and parameters used in this work were chosen to minimize any detrimental effects on these two important outcomes, as in-

vestigated in this chapter. The analysis presented in this chapter demonstrates the effectiveness and robustness of the developed techniques, and serves to identify the potential pitfalls that may arise during the implementation of such automated data processing techniques.

In chapter 5, it is shown that the inherent variability in Raman spectra of single DU145 cells cultured *in vitro* is correlated with biochemical changes arising from (1) cell cycle progression and (2) the changing confluency of the cell culture during the first 3-4 days after sub-culturing. The variability between single-cell Raman spectra arising from cell cycle progression is shown to arise from varying intensities of protein and nucleic acid features relative to lipid features. Raman spectra acquired from synchronized cell cultures demonstrate a continual increase in the average nucleic acid and protein content relative to lipid content as cells progress from early G1 phase to the G1/S boundary and into S phase. The molecular origins of the Raman features affected by cell cycle progression are identified, for both the LWN and the HWN spectral windows, by the features of the first PCA component for each window (figures 5.3 and 5.11). Application of PCA demonstrates that in the absence of additional external sources of variability, cell cycle variability typically accounts for 40-60% of the total variance if the LWN window is used and 75-90% if the HWN window is used. The characterization of cell cycle variability presented in this chapter is useful for future Raman studies in order to distinguish the inherent cell cycle variability between cells from other independent sources of variability (*e.g.*, radiation induced responses). The molecular origins of the Raman features producing variability correlated with the changing confluency of a DU145 cell culture are identified for the LWN spectral window by the features of the second PCA component from the study of asynchronous cell cultures (figure 5.6a). In this work with DU145 cells, this source of variability explains up to 17% of the total variance if the LWN window is used. The characterization of this variability, as presented in this chapter, is an important consideration for future Raman studies involving comparisons between cell cultures harvested at different time intervals after sub-culturing. For example, if cell cultures are allowed to incubate after sub-culturing for 3-4 days before Raman analysis, it is shown that the variance explained by this source of variability is greatly reduced. In chapter 6, this reduction in spectral variability is shown to facilitate a clearer observation of other spectral differences between cell cultures arising from a radiation induced response. Overall, the results presented in chapter 5 demonstrate the levels of sensitivity and specificity achieved by the RS acquisition and data processing

techniques developed in this work for analysis of biochemical changes in single human cells.

The primary DU145 irradiation study presented in chapter 6 is the first use of RS to observe radiation induced biochemical changes in single human cells cultured *in vitro*. The sensitivity of RS to different types of biomolecules in a single acquisition is shown to enable the noninvasive and nondestructive observation of radiation induced effects in single living cells. It is demonstrated that the application of PCA effectively discriminates between cell cycle related spectral changes and the spectral changes arising from a radiation induced biochemical response. The radiation induced spectral changes observed in irradiated DU145 cells are found to be isolated to the second PCA component obtained from the data set and are assigned to various nucleic acids, lipids, amino acids and conformational protein structures (table 6.1). Analysis of the PCA scores for the second PCA component indicates that the magnitude of the spectral changes observed increase with both the irradiated dose (0 to 50 Gy) and the post-irradiation incubation time (0 to 120 hours). This successful demonstration of the sensitivity of RS to radiation induced changes in single cells demonstrates that RS is indeed an attractive modality for radiobiological studies of living biological systems, and provides the foundation for subsequent studies involving different *in vitro* cell lines and different models of cancer, presented in chapter 7.

In the first study presented in chapter 7, a preliminary panel of seven human tumour cell lines derived from prostate (DU145, PC3, LNCaP and PacMet), breast (MDA-MB-231 and MCF7) and lung (H460) is used to demonstrate that RS can detect biochemical signatures of *in vitro* radiation response (*i.e.*, radiation induced PCA components) that segregate according to p53 status and intrinsic radiosensitivity (SF_2). Through analysis of the corresponding PCA scores, the observed RS signatures are found to arise from radiation induced changes in the cellular concentrations of aromatic amino acids, conformational protein structures, and certain nucleic acid and lipid functional groups, and are detected from live cells analyzed 1-3 days post-irradiation. This demonstrated sensitivity to the biomolecules responsible for the observed radiation responses provides new insight into possible mechanisms of radiation survival, and into the differences in such survival mechanisms between wt p53 and mt p53 cell lines. Potential radiation induced biochemical response mechanisms underlying these RS observations, which may increase or decrease tumour cell survival post-irradiation, are postulated. Preliminary results obtained from a mouse prostate tumour cell line (TRAMP-C2), irradiated both *in vitro* and *in vivo* (as solid tumours

in mice), indicate that RS signatures of radiation resistance or sensitivity observed *in vitro* may also be detectable from tumour cells obtained from an irradiated solid tumour.

Implications for future work

The results presented in chapters 6 and 7 demonstrate the utility of using RS for radiobiological investigations at the single-cell level. Specifically, the observed relationship between the RS radiation response signatures and intrinsic radiosensitivity supports the possibility of using RS for detecting radiation resistance or sensitivity in clinical practice. Future work may lead to the development of RS techniques for monitoring or predicting tumour response in radiation therapy patients. Although the single fraction doses of radiation used in this work are very high from a clinical standpoint ($15 - 50$ Gy versus $1 - 2.5$ Gy per fraction), there are other established techniques for enhancing the sensitivity of biological Raman signals, such as surface enhanced Raman scattering using targeted nanoprobes [213–217], which may prove useful for detecting radiation responses with RS at lower, more clinically relevant doses.

The results of the first study in chapter 7 suggest that certain biochemical changes across proteins, nucleic acids and lipids, observed with RS noninvasively from live, unfixed, single cells 1-3 days post-irradiation, may be the result of cellular mechanisms that promote cell survival after radiation exposure. If confirmed, this result would make RS a very attractive technique for future studies designed to understand, monitor or predict radiation response. However, future work must initially extend these results to a larger panel of tumour cell lines to further investigate the correlation between observed RS biochemical radiation response and clonogenic radiation survival *in vitro*. Clonogenic measurements of intrinsic radiosensitivity performed by the same laboratory may provide improved radiosensitivity discrimination across the tested cell lines, possibly allowing for improved comparisons and correlations with RS observations. In addition, complementary biochemical analysis techniques with superior molecular specificity as compared to RS, such as flow cytometry or other conventional staining assays, may be helpful in elucidating a biological explanation for the radiation response signatures observed with RS (*e.g.*, radiation induced expression of known survival signals or anti-apoptosis factors).

It is also necessary for future RS studies to investigate the biochemical radiation response of human tumour cells grown and irradiated as solid tumours in mice, with

comparable radiation doses as used in this work. The second study presented in chapter 7 demonstrates that the primary sources of spectral variability observed from a mouse prostate tumour cell line (TRAMP-C2) irradiated *in vitro* are also observed when the same cell line is grown and irradiated as a solid tumour *in vivo* prior to extraction of tumour cells for RS analysis. This preliminary result suggests that the candidate RS signatures of radiation resistance and sensitivity observed in this work *in vitro* may also be detectable from tumour cells obtained from patients undergoing radiation therapy treatment. It is encouraging to note that the *in vitro* radiation response of the TRAMP-C2 cell line is the weakest of the *in vitro* responses observed, and yet a similar radiation "enhanced" response is still observed from TRAMP-C2 cells grown and irradiated *in vivo*. As such, it is likely that the inherently stronger radiation responses of the human tumour cell lines investigated in chapter 7 may also be observed with RS from tumour cells grown and irradiated as solid tumours in mice. Modifications to the sample preparation and RS acquisition techniques, as suggested in section 7.6.4, will improve the robustness and portability of the RS techniques for such future studies.

It is also postulated in chapter 7 (sections 7.4.4 and 7.6.2) that the radiation responses observed with RS in this work may not be unique to radiation exposure. Both the *in vitro* radiation response signature obtained from the H460 cell line and the candidate *in vitro* and *in vitro* radiation response signatures obtained from the TRAMP-C2 cell line are also obtained, with reduced percent variance explained, from data sets composed of only unirradiated tumour cells. This result may indicate that other factors, such as stresses imposed on tumour cells by the tumour microenvironment, may induce a biochemical response in tumour cells similar to the response observed post-irradiation. If confirmed, it is possible that RS analysis of such biochemical sources of variability in tumour cells prior to irradiation may have merit for prediction of radiation resistance or sensitivity in clinical radiation therapy treatments.

Although this work has focused on investigating the response of tumour cells to radiation therapy, the methods developed may also be applied to investigate tumour cell response to other forms of cytotoxic cancer therapy, such as chemotherapy. Investigating biochemical mechanisms of chemotherapy drug response, including possible mechanisms of resistance or sensitivity, may also provide significant benefit to the advancement of personalized cancer therapy.

Bibliography

[1] WC Röntgen. On a New Kind of Rays. *Nature*, 53(1369):274–276, 1896.

[2] JE Aldrich and BC Lentle. A New Kind of Ray: The Radiological Sciences in Canada. *University of British Columbia, Vancouver, BC*, 1995.

[3] D Brian. The Curies: A Biography of the Most Controversial Family in Science. *J. Wiley, Hoboken, NJ*, 2005.

[4] DI Thwaites and JB Tuohy. Back to the future: the history and development of the clinical linear accelerator. *Phys. Med. Biol.*, 51:R343–R362, 2006.

[5] EM Soffen, GE Hanks, CC Hwang, and JC Chu. Conformal static field therapy for low volume low grade prostate cancer with rigid immobilization. *Int. J. Radiat. Oncol. Biol. Phys.*, 20(1):141–6, 1991.

[6] DJ Convery and ME Rosenbloom. The generation of intensity-modulated fields for conformal radiotherapy by dynamic collimation. *Phys. Med. Biol.*, 37(6):1359–1374, 1992.

[7] TR Mackie, J Kapatoes, K Ruchala, W Lu, C Wu, G Olivera, L Forrest, W Tome, J Welsh, R Jeraj, P Harari, P Reckwerdt, B Paliwal, M Ritter, H Keller, J Fowler, and M Mehta. Image guidance for precise conformal radiotherapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 56(1):89–105, 2003.

[8] TR Mackie, T Holmes, S Swerdloff, P Reckwerdt, JO Deasy, J Yang, B Paliwal, and T Kinsella. Tomotherapy: a new concept for the delivery of dynamic conformal radiotherapy. *Med. Phys.*, 20(6):1709–19, 1993.

[9] K Otto. Volumetric modulated arc therapy: IMRT in a single gantry arc. *Med. Phys.*, 35(1):310–7, 2008.

[10] FM Khan. The Physics of Radiation Therapy, 3rd edition. *Lippincott Williams & Wilkins, Philadelphia, PA*, 2003.

[11] NCIC. Canadian Cancer Statistics. *National Cancer Institute of Canada*, 2011.

[12] P Rubin. Clinical Oncology, 8th edition. *W. B. Saunders Company, Philadelphia, PA*, 2001.

[13] S Bentzen and J Overgaard. Patient-to-patient variability in the expression of radiation-induced normal tissue injury. *Semin. Radiat. Oncol.*, 4(2):68–80, 1994.

[14] HE Johns and JR Cunningham. The Physics of Radiology. *Charles C. Thomas Publisher, Springfield, Il*, 1983.

[15] EL Travis. Medical Radiobiology, 2nd edition. *Year Book Medical Publishers, Inc., Chicago, Il*, 1989.

[16] WA Hendrickson and MM Teeter. Structure of the hydrophobic protein crambin determined directly from the anomalous scattering of sulfur. *Nature*, 290(5802):107–113, 1981.

[17] HA Crissman, Z Darzynkiewicz, RA Tobey, and JA Steinkamp. Correlated measurements of DNA, RNA, and protein in individual cells by flow-cytometry. *Science*, 228(4705):1321–1324, 1985.

[18] Z Darzynkiewicz, T Sharpless, L Staianocioco, and MR Melamed. Subcompartments of the G1 phase of cell-cycle detected by flow cytometry. *P. Natl. Acad. Sci. USA*, 77(11):6696–6699, 1980.

[19] JW Gray, F Dolbeare, MG Pallavicini, W Beisker, and F Waldman. Cell-cycle analysis using flow-cytometry. *Int. J. Radiat. Biol.*, 49:237–255, 1986.

[20] EJ Hall and AJ Giaccia. Radiobiology for the Radiologist, 6th edition. *Lippincott Williams & Wilkins, Philadelphia, PA*, 2006.

[21] Franklin Hutchinson. The Distance That a Radical Formed by Ionizing Radiation Can Diffuse in a Yeast Cell. *Radiat. Res.*, pages 473–483, 1957.

[22] M Dizdaroglu. Application of capillary gas-chromatography mass-spectrometry to chemical characterization of radiation-induced base damage of DNA - Implications for assessing DNA-repair processes. *Anal Biochem*, 144:593–603, 1985.

[23] DD Ager, WC Dewey, K Gardiner, W Harvey, RT Johnson, and CA Waldren. Measurement of radiation-induced DNA double-strand breaks by pulsed-field gel-electrophoresis. *Radiat. Res.*, 122:181–187, 1990.

[24] M Bergqvist, D Brattstrom, M Stalberg, H Vaghef, O Brodin, and B Hellman. Evaluation of radiation-induced DNA damage and DNA repair in human lung cancer cell lines with different radiosensitivity using alkaline and neutral single cell gel electrophoresis. *Cancer Lett*, 133:9–18, 1998.

[25] VS Dhillon, P Thomas, and M Fenech. Comparison of DNA damage and repair following radiation challenge in buccal cells and lymphocytes using single-cell gel electrophoresis. *Int. J. Radiat. Biol.*, 80:517–528, 2004.

[26] Z Cai, Z Chen, KE Bailey, DA Scollard, RM Reilly, and KA Vallis. Relationship between induction of phosphorylated H2AX and survival in breast cancer cells exposed to 111In-DTPA-hEGF. *J. Nucl. Med.*, 49(8):1353–61, 2008.

[27] NAP Franken, HM Rodermond, J Stap, J Haveman, and C van Bree. Clonogenic assay of cells in vitro. *Nat. Protoc.*, 1(5):2315–9, 2006.

[28] Q Matthews. Development of a Raman microscope for applications in radiobiology. *M.Sc. Thesis: University of Victoria*, 2008.

[29] NCRP. Report #150: Extrapolation of radiation-induced cancer risks from non-human experimental systems to humans. *National Council on Radiation Protection & Measurements*, 2005.

[30] LJ Peters. Radiation therapy tolerance limits - For one or for all? Janeway lecture. *Cancer*, 77:2379–2385, 1996.

[31] M Joiner and A van der Kogel. Basic Clinical Radiobiology, 4th edition. *Hodder Arnold, London, UK*, 2009.

[32] FB Geara, LJ Peters, KK Ang, JL Wike, and WA Brock. Prospective comparison of in-vitro normal-cell radiosensitivity and normal tissue-reactions in radiotherapy patients. *Int. J. Radiat. Oncol.*, 27(5):1173–1179, 1993.

[33] WA Brock and SL Tucker. In vitro radiosensitivity and normal tissue damage. *Radiother. Oncol.*, 55(2):93–94, 2000.

[34] E López, R Guerrero, MI Nunez, R del Moral, M Villalobos, J Martinez-Galan, MT Valenzuela, J Munoz-Gamez, FJ Oliver, D Martin-Oliva, and JMR de Almodovar. Early and late skin reactions to radiotherapy for breast cancer and their correlation with radiation-induced DNA damage in lymphocytes. *Breast Cancer Res.*, 7:R690–R698, 2005.

[35] CN Andreassen and J Alsner. Genetic variants and normal tissue toxicity after radiotherapy: A systematic review. *Radiother. Oncol.*, 92(3):299–309, 2009.

[36] JM Brown and AC Koong. High-dose single-fraction radiotherapy: exploiting a new biology? *Int. J. Radiat. Oncol. Biol. Phys.*, 71(2):324–5, 2008.

[37] S Tapio and V Jacob. Radioadaptive response revisited. *Radiat. Environ. Biophys.*, 46(1):1–12, 2007.

[38] H Nagasawa and JB Little. Induction of sister chromatid exchanges by extremely low-doses of α-particles. *Cancer Res.*, 52(22):6394–6396, 1992.

[39] M Folkard, KM Prise, AG Michette, and B Vojnovic. The use of radiation microbeams to investigate the bystander effect in cells and tissues. *Nucl. Instrum. Meth. A*, 580(1):446–450, 2007.

[40] WF Morgan and MB Sowa. Non-targeted bystander effects induced by ionizing radiation. *Mutat. Res.*, 616(1-2):159–164, 2007.

[41] EL Levine, A Renehan, R Gossiel, SE Davidson, SA Roberts, C Chadwick, DP Wilks, CS Potten, JH Hendry, RD Hunter, and CML West. Apoptosis, intrinsic radiosensitivity and prediction of radiotherapy response in cervical-carcinoma. *Radiother. Oncol.*, 37(1):1–9, 1995.

[42] CML West, SE Davidson, SA Roberts, and RD Hunter. The independence of intrinsic radiosensitivity as a prognostic factor for patient response to radiotherapy of carcinoma of the cervix. *Brit. J. Cancer*, 76(9):1184–1190, 1997.

[43] T Björk-Eriksson, C West, E Karlsson, and C Mercke. Tumor radiosensitivity (SF2) is a prognostic factor for local control in head and neck cancers. *Int. J. Radiat. Oncol. Biol. Phys.*, 46(1):13–9, 2000.

[44] M Nordsmark and J Overgaard. A confirmatory prognostic study on oxygenation status and loco-regional control in advanced head and neck squamous cell carcinoma treated by radiation therapy. *Radiother. Oncol.*, 57(1):39–43, 2000.

[45] M Luukkaa, T Jokilehto, P Kronqvist, T Vahlberg, R Grenman, P Jaakkola, and H Minn. Expression of the cellular oxygen sensor PHD2 (EGLN-1) predicts radiation sensitivity in squamous cell cancer of the head and neck. *Int. J. Radiat. Biol.*, 85(10):900–908, 2009.

[46] P Vaupel and A Mayer. Hypoxia in cancer: significance and impact on clinical outcome. *Cancer Metast. Rev.*, 26(2):225–239, 2007.

[47] AC Begg, K Haustermans, AAM Hart, S Dische, M Saunders, B Zackrisson, H Gustaffson, P Coucke, N Paschoud, M Hoyer, J Overgaard, P Antognoni, A Richetti, J Bourhis, H Bartelink, JC Horiot, R Corvo, W Giaretti, H Awwad, T Shouman, T Jouffroy, Z Maciorowski, W Dobrowsky, H Struikmans, D Rutgers, and GD Wilson. The value of pretreatment cell kinetic parameters as predictors for radiotherapy outcome in head and neck cancer: a multicenter analysis. *Radiother. Oncol.*, 50(1):13–23, 1999.

[48] AC Begg. Predicting response to radiotherapy: Evolutions and revolutions. *Int. J. Radiat. Biol.*, 85(10):825–36, 2009.

[49] AJ McIlwrath, PA Vasey, GM Ross, and R Brown. Cell-cycle arrests and radiosensitivity of human tumor-cell lines: Dependence on wild-type p53 for radiosensitivity. *Cancer Res*, 54(14):3718–3722, 1994.

[50] N Oya, F Zolzer, F Werner, and C Streffer. Effects of serum starvation on radiosensitivity, proliferation and apoptosis in four human tumor cell lines with different p53 status. *Strahlenther. Onkol.*, 179:99–106, 2003.

[51] SL Scott, JD Earle, and PH Gumerlock. Functional p53 increases prostate cancer cell survival after exposure to fractionated doses of ionizing radiation. *Cancer Res.*, 63(21):7190–6, 2003.

[52] JR Williams, Y Zhang, J Russell, C Koch, and JB Little. Human tumor cells segregate into radiosensitivity groups that associate with ATM and TP53 status. *Acta Oncol.*, 46(5):628–38, 2007.

[53] JR Williams, Y Zhang, H Zhou, DS Gridley, CJ Koch, J Russell, JS Slater, and JB Little. A quantitative overview of radiosensitivity of human tumor cells across histological type and TP53 status. *Int. J. Radiat. Biol.*, 84(4):253–64, 2008.

[54] JF Torres-Roca, S Eschrich, HY Zhao, G Bloom, J Sung, S McCarthy, AB Cantor, A Scuto, CG Li, SM Zhang, R Jove, and T Yeatman. Prediction of radiation sensitivity using a gene expression classifier. *Cancer Res.*, 65(16):7169–7176, 2005.

[55] S Eschrich, H Zhang, H Zhao, D Boulware, JH Lee, G Bloom, and JF Torres-Roca. Systems biology modeling of the radiation sensitivity network: a biomarker discovery platform. *Int. J. Radiat. Oncol. Biol. Phys.*, 75(2):497–505, 2009.

[56] SA Eschrich, J Pramana, H Zhang, H Zhao, D Boulware, J Lee, G Bloom, C Rocha-Lima, S Kelley, DP Calvin, TJ Yeatman, AC Begg, and JF Torres-Roca. A gene expression model of intrinsic tumor radiosensitivity: Prediction of response and prognosis after chemoradiation. *Int. J. Radiat. Oncol. Biol. Phys.*, 75(2):489–496, 2009.

[57] GJ Puppels, JHF Olminkhof, GMJ Segersnolten, C Otto, FFM Demul, and J Greve. Laser irradiation and Raman spectroscopy of single living cells and chromosomes: sample degradation occurs with 514.5 nm but not with 660 nm laser light. *Exp. Cell Res.*, 195(2):361–367, 1991.

[58] I Notingher, S Verrier, H Romanska, AE Bishop, JM Polak, and LL Hench. In situ characterisation of living cells by Raman spectroscopy. *Spectrosc.-Int. J.*, 16(2):43–51, 2002.

[59] I Notingher, S Verrier, S Haque, JM Polak, and LL Hench. Spectroscopic study of human lung epithelial cells (A549) in culture: Living cells versus dead cells. *Biopolymers*, 72:230–240, 2003.

[60] JR Mourant, J Dominguez, S Carpenter, KW Short, TM Powers, R Michalczyk, N Kunapareddy, A Guerra, and JP Freyer. Comparison of vibrational spectroscopy to biochemical and flow cytometry methods for analysis of the basic biochemical composition of mammalian cells. *J. Biomed. Opt.*, 11(6):064024, 2006.

[61] SP Verma and N Sonwalkar. Structural-changes in plasma-membranes prepared from irradiated Chinese-hamster V79-cells as revealed by Raman-spectroscopy. *Radiat. Res.*, 126(1):27–35, 1991.

[62] JT Motz, M Hunter, LH Galindo, JA Gardecki, JR Kramer, RR Dasari, and MS Feld. Optical fiber probe for biomedical Raman spectroscopy. *Appl. Optics.*, 43(3):542–54, 2004.

[63] GJ Puppels, FFM Demul, C Otto, J Greve, M Robertnicoud, DJ Arndtjovin, and TM Jovin. Studying single living cells and chromosomes by confocal Raman microspectroscopy. *Nature*, 347(6290):301–303, 1990.

[64] D Naumann, S Keller, D Helm, C Schulz, and B Schrader. FT-IR Spectroscopy and FT-Raman Spectroscopy are powerful analytical tools for the noninvasive characterization of intact microbial-cells. *J. Mol. Struct.*, 347:399–405, 1995.

[65] KC Schuster, I Reese, E Urlaub, JR Gapes, and B Lendl. Multidimensional information on the chemical composition of single bacterial cells by confocal Raman microspectroscopy. *Anal. Chem.*, 72:5529–5534, 2000.

[66] KC Schuster, E Urlaub, and JR Gapes. Single-cell analysis of bacteria by Raman microscopy: spectral information on the chemical composition of cells and on the heterogeneity in a culture. *J. Microbiol. Meth.*, 42:29–38, 2000.

[67] K Maquelin, C Kirschner, LP Choo-Smith, N van den Braak, HP Endtz, D Naumann, and GJ Puppels. Identification of medically relevant microorganisms by vibrational spectroscopy. *J. Microbiol. Meth.*, 51:255–271, 2002.

[68] A Nijssen, TCB Schut, F Heule, PJ Caspers, DP Hayes, MHA Neumann, and GJ Puppels. Discriminating basal cell carcinoma from its surrounding tissue by Raman spectroscopy. *J. Invest. Dermatol.*, 119(1):64–69, 2002.

[69] J Choi, J Choo, H Chung, DG Gweon, J Park, HJ Kim, S Park, and CH Oh. Direct observation of spectral differences between normal and basal cell carcinoma (BCC) tissues using confocal Raman microscopy. *Biopolymers*, 77:264–272, 2005.

[70] C Lieber, S Majumder, D Billheimer, D Ellis, and A Mahadevan-Jansen. Raman microspectroscopy for skin cancer detection in vitro. *J. Biomed. Opt.*, 13(2):024013, 2008.

[71] BWD de Jong, TCS Bakker, K Maquelin, T van der Kwast, CH Bangma, DJ Kok, and GJ Puppels. Discrimination between nontumor bladder tissue and tumor by Raman spectroscopy. *Anal. Chem.*, 78(22):7761–7769, 2006.

[72] SK Teh, W Zheng, KY Ho, M Teh, KG Yeoh, and Z Huang. Diagnosis of gastric cancer using near-infrared Raman spectroscopy and classification and regression tree techniques. *J. Biomed. Opt.*, 13(3):034013, 2008.

[73] S Teh, W Zheng, K Ho, M Teh, and K Yeoh. Near-infrared Raman spectroscopy for gastric precancer diagnosis. *J. Raman Spectrosc.*, 40:908–914, 2009.

[74] M Tollefson, J Magera, T Sebo, J Cohen, A Drauch, J Maier, and I Frank. Raman spectral imaging of prostate cancer: can Raman molecular imaging be used to augment standard histopathology? *BJU International*, 106(4):484–488, 2010.

[75] KM Omberg, JC Osborn, SLL Zhang, JP Freyer, JR Mourant, and JR Schoonover. Raman spectroscopy and factor analysis of tumorigenic and non-tumorigenic cells. *Appl. Spectrosc.*, 56(7):813–819, 2002.

[76] L Notingher, G Jell, PL Notingher, I Bisson, O Tsigkou, JM Polak, MM Stevens, and LL Hench. Multivariate analysis of Raman spectra for in vitro non-invasive studies of living cells. *J. Mol. Struct.*, 744:179–185, 2005.

[77] N Stone, C Kendall, N Shepherd, P Crow, and H Barr. Near-infrared Raman spectroscopy for the classification of epithelial pre-cancers and cancers. *J. Raman Spectrosc.*, 33(7):564–573, 2002.

[78] CM Krishna, GD Sockalingum, G Kegelaer, S Rubin, VB Kartha, and M Manfait. Micro-Raman spectroscopy of mixed cancer cell populations. *Vib. Spectrosc.*, 38:95–100, 2005.

[79] P Crow, B Barrass, C Kendall, M Hart-Prieto, M Wright, R Persad, and N Stone. The use of Raman spectroscopy to differentiate between different prostatic adenocarcinoma cell lines. *Brit. J. Cancer.*, 92:2166–2170, 2005.

[80] TJ Harvey, EC Faria, A Henderson, E Gazi, AD Ward, NW Clarke, MD Brown, RD Snook, and P Gardner. Spectral discrimination of live prostate and bladder cancer cell lines using Raman optical tweezers. *J. Biomed. Opt.*, 13(6):064004, 2008.

[81] S Verrier, I Notingher, JM Polak, and LL Hench. In situ monitoring of cell death using Raman microspectroscopy. *Biopolymers*, 74:157–162, 2004.

[82] N Kunapareddy, J Freyer, and J Mourant. Raman spectroscopic characterization of necrotic cell death. *J. Biomed. Opt.*, 13(5):054002–1 – 054002–9, 2008.

[83] KW Short, S Carpenter, JP Freyer, and JR Mourant. Raman spectroscopy detects biochemical changes due to proliferation in mammalian cell cultures. *Biophys. J.*, 88(6):4274–4288, 2005.

[84] RJ Swain, G Jell, and MA Stevens. Non-invasive analysis of cell cycle dynamics in single living cells with Raman micro-spectroscopy. *J. Cell Biochem.*, 104(4):1427–1438, 2008.

[85] K Sailer, S Viaggi, and M Nusse. Radiation-induced structural modifications in dsDNA analysed by FT-Raman spectroscopy. *Int. J. Radiat. Biol.*, 69:601–613, 1996.

[86] CP Shaw and A Jirasek. The Use of Ultraviolet Resonance Raman Spectroscopy in the Analysis of Ionizing-Radiation-Induced Damage in DNA. *Appl. Spectrosc.*, 63(4):412–22, 2009.

[87] A Synytsya, A Synytsya, P Alexa, R Wagner, M Davídková, and K Volka. Raman spectroscopic study on sodium hyaluronate: an effect of proton and γ irradiation. *J. Raman Spectrosc.*, 42(1):544–550, 2011.

[88] SP Verma. Low-levels of irradiation modify lipid domains in model membranes - a laser Raman-study. *Radiat. Res.*, 107(2):183–193, 1986.

[89] SP Verma and A Rastogi. Role of proteins in protection against radiation-induced damage in membranes. *Radiat. Res.*, 122(2):130–136, 1990.

[90] SP Verma, A Singhal, and N Sonwalkar. Ionizing-radiation target groups of band-3 inserted into egg lecithin liposomes as determined by Raman-spectroscopy. *Int. J. Radiat. Biol.*, 63(3):279–288, 1993.

[91] K Sailer, S Viaggi, and M Nüsse. Kinetics of radiation- and cytochrome c-induced modifications in liposomes analysed by FT-Raman spectroscopy. *Biochim. Biophys. Acta*, 1329(2):259–68, 1997.

[92] A Synytsya, P Alexa, J Besserer, J De Boer, S Froschauer, R Gerlach, M Loewe, M Moosburger, I Obstova, P Quicken, B Sosna, K Volka, and M Wurkner. Raman spectroscopy of tissue samples irradiated by protons. *Int. J. Radiat. Biol.*, 80(8):581–591, 2004.

[93] RJ Lakshmi, VB Kartha, CM Krishna, JGR Solomon, G Ullas, and PU Devi. Tissue Raman spectroscopy for the study of radiation damage: Brain irradiation of mice. *Radiat. Res.*, 157:175–182, 2002.

[94] MS Vidyasagar, K Maheedhar, BM Vadhiraja, DJ Fernendes, VB Kartha, and CM Krishna. Prediction of radiotherapy response in cervix cancer by Raman spectroscopy: A pilot study. *Biopolymers*, 89(6):530–537, 2008.

[95] A Smekal. Zur Quantentheorie der Dispersion. *Naturwissenschaften II*, 43:873–875, 1923.

[96] CV Raman and KS Krishnan. A new type of secondary radiation. *Nature*, 121:501–502, 1928.

[97] J Brand. The Discovery of the Raman Effect. *Notes and Records of the Royal Society of London*, pages 1–23, 1989.

[98] CV Raman and KS Krishnan. A new class of spectra due to secondary radiation. *Indian J. Phys.*, 2:399–419, 1928.

[99] FP Kerschbaum. *Z. Instrumentenk*, 34:43, 1914.

[100] B Veskatesachar and L Sibaiya. *Indian J. Phys.*, 5:747, 1930.

[101] JH Hibben. The Raman Effect and Its Chemical Applications. *Reinhold Publishing Corp., New York, NY*, 1939.

[102] TR Gilson and PJ Hendra. Laser Raman Spectroscopy. *Wiley-Interscience, London, UK*, 1970.

[103] DH Rank, RJ Pfister, and PD Coleman. Photoelectric detection and intensity measurement in Raman spectra. *J. Opt. Soc. Am.*, 32:390–396, 1942.

[104] RF Stamm, CF Salzman, and T Mariner. Photoelectric Raman spectrometer with automatic range changing. I. Conversion of photographic instrument. *J. Opt. Soc. Am.*, 43(2):119–125, 1953.

[105] Y Talmi. Spectrophotometry and spectrofluorometry with the self-scanned photo-diode array. *Appl. Spectrosc.*, 36(1):1–18, 1982.

[106] DG Jones. Photodiode array detectors in UV-VIS spectroscopy. Part 1. *Anal. Chem.*, 57(9):1057A, 1985.

[107] PM Epperson, JV Sweedler, RB Bilhorn, GR Sims, and MB Denton. Applications of charge-transfer devices in spectroscopy. *Anal. Chem.*, 60(5):327A–335A, 1988.

[108] JE Pemberton, RL Sobocinski, and GR Sims. The effect of charge traps on Raman-spectroscopy using a Thomson-CSF charge coupled device detector. *Appl. Spectrosc.*, 44(2):328–330, 1990.

[109] E Smith and G Dent. Modern Raman Spectroscopy: A Practical Approach. *John Wiley & Sons Ltd, Chichester, England*, 2005.

[110] JR Ferraro and K Nakamoto. Introductory Raman Spectroscopy. *Academic Press, Inc., San Diego, CA*, 1994.

[111] J Tang and AC Albrecht. Developments in the theories of vibrational Raman intensities, *in* "Raman Spectroscopy" (HA Szymanski, Ed.) Vol. 2. *Plenum Press, New York, NY*, 1970.

[112] RL McCreery. Raman Spectroscopy for Chemical Analysis. *John Wiley & Sons, Inc., New York, NY*, 2000.

[113] JM Lerner and A Thevenon. The Optics of Spectroscopy - A Tutorial. *HORIBA Jobin Yvon Inc., Edison, NJ*, 1988.

[114] C Dyer and BJE Smith. Application of continuous extended scanning techniques to the simultaneous detection of Raman-scattering and photoluminescence from calcium disilicates using visible and near-infrared excitation. *J. Raman Spectrosc.*, 26(8-9):777–785, 1995.

[115] KR Spring and MW Davidson. Basic Concepts and Formulas in Microscopy. *Molecular Expressions, Tallahassee, FL*, 2008.

[116] JM Lerner and A Thevenon. An Introduction to Raman Spectroscopy. *HORIBA Jobin Yvon Inc., Edison, NJ*, 1988.

[117] I Notingher, JR Jones, S Verrier, I Bisson, P Embanga, P Edwards, JM Polak, and LL Hench. Application of FTIR and Raman spectroscopy to characterisation of bioactive materials and living cells. *Spectrosc.-Int. J.*, 17(2-3):275–288, 2003.

[118] I Notingher and LL Hench. Raman microspectroscopy: a noninvasive tool for studies of individual living cells in vitro. *Expert Rev. Med. Devic.*, 3(2):215–234, 2006.

[119] C Krafft, T Knetschke, A Siegner, RHW Funk, and R Salzer. Mapping of single cells by near infrared Raman microspectroscopy. *Vib. Spectrosc.*, 32:75–83, 2003.

[120] N Uzunbajakava, A Lenferink, Y Kraan, B Willekens, G Vrensen, J Greve, and C Otto. Nonresonant Raman imaging of protein distribution in single human cells. *Biopolymers*, 72:1–9, 2003.

[121] D Borchman, DX Tang, and MC Yappert. Lipid composition, membrane structure relationships in lens and muscle sarcoplasmic reticulum membranes. *Biospectroscopy*, 5(3):151–167, 1999.

[122] G Schulze, A Jirasek, MML Yu, A Lim, RFB Turner, and MW Blades. Investigation of selected baseline removal techniques as candidates for automated implementation. *Appl. Spectrosc.*, 59:545–574, 2005.

[123] JR Mourant, KW Short, S Carpenter, N Kunapareddy, L Coburn, TM Powers, and JP Freyer. Biochemical differences in tumorigenic and nontumorigenic cells measured by Raman and infrared spectroscopy. *J. Biomed. Opt.*, 10(3):031106, 2005.

[124] C Plathow and WA Weber. Tumor cell metabolism imaging. *J. Nucl. Med.*, 49 Suppl 2:43S–63S, 2008.

[125] P Laverman, OC Boerman, FHM Corstens, and WJG Oyen. Fluorinated amino acids for tumour imaging with positron emission tomography. *Eur. J. Nucl. Med.*, 29(5):681–90, 2002.

[126] S Gross and D Piwnica-Worms. Spying on cancer: molecular imaging in vivo with genetically encoded reporters. *Cancer Cell*, 7(1):5–15, 2005.

[127] K Shibuya, E Yoshida, F Nishikido, T Suzuki, T Tsuda, N Inadama, T Yamaya, and H Murayama. Limit of Spatial Resolution in FDG-PET due to Annihilation Photon Non-Collinearity. In R Magjarevic, R Magjarevic, and JH Nagel, editors, *World Congress on Medical Physics and Biomedical Engineering 2006*, volume 14 of *IFMBE Proceedings*, pages 1667–1671. Springer Berlin Heidelberg.

[128] MA Thomas, T Lange, SS Velan, R Nagarajan, S Raman, A Gomez, D Margolis, S Swart, RR Raylman, RF Schulte, and P Boesiger. Two-dimensional MR spectroscopy of healthy and cancerous prostates in vivo. *Magn. Reson. Mater. Phy.*, 21(6):443–58, 2008.

[129] F Estève, C Rubin, S Grand, H Kolodié, and JF Le Bas. Transient metabolic changes observed with proton MR spectroscopy in normal human brain after radiation therapy. *Int. J. Radiat. Oncol. Biol. Phys.*, 40(2):279–86, 1998.

[130] AR Khan, P Rana, MM Devi, S Chaturvedi, S Javed, RP Tripathi, and S Khushu. Nuclear magnetic resonance spectroscopy-based metabonomic investigation of biochemical effects in serum of γ-irradiated mice. *Int. J. Radiat. Biol.*, 87(1):91–7, 2011.

[131] R Tarnawski, M Sokol, P Pieniazek, B Maciejewski, J Walecki, L Miszczyk, and T Krupska. 1H-MRS in vivo predicts the early treatment outcome of postoperative radiotherapy for malignant gliomas. *Int. J. Radiat. Oncol. Biol. Phys.*, 52(5):1271–6, 2002.

[132] GM Clore and AM Gronenborn. Determination of three-dimensional structures of proteins and nucleic acids in solution by nuclear magnetic resonance spectroscopy. *Crit. Rev. Biochem. Mol. Biol.*, 24(5):479–564, 1989.

[133] S Grande, AM Luciani, A Rosi, R Cherubini, M Conzato, L Guidoni, and V Viti. Radiation effects on soluble metabolites in cultured HeLa cells examined by 1H MRS: changes in concentration of glutathione and of lipid catabolites induced by gamma rays and proton beams. *Int. J. Cancer.*, 96 (Suppl.):27–42, 2001.

[134] M Yoshikawa and N Nagai. *in* "Handbook of Vibrational Spectroscopy" (J Chalmers and P Griffiths, Eds.) Vol. 4. *John Wiley & Sons, Inc., New York, NY*, pages 2593–2600, 2001.

[135] Y Nishimura and M Tsuboi. Local conformations and polymorphism of DNA duplexes as revealed by their Raman spectra, *in* "Advances in Spectroscopy" (RJH Clark and RE Hester, Eds.) Vol. 13. *John Wiley, New York, NY*, 1986.

[136] P Dhamelincourt, F Wallart, M Leclercq, AT N'Guyen, and DO Landon. Laser Raman molecular microprobe (MOLE). *Anal. Chem.*, 51:414A–421A, 1979.

[137] A Paipetis, C Vlattas, and C Galiotis. Remote laser Raman microscopy (ReRaM). 1 - Design and testing of a confocal microprobe. *J. Raman Spectrosc.*, 27(7):519–526, 1996.

[138] J Chalmers and P Griffiths (Eds.). Handbook of Vibrational Spectroscopy, Vol. 4. *John Wiley & Sons, Inc., New York, NY*, 2001.

[139] MD Schaeberle, DD Tuschel, and PJ Treado. Raman chemical imaging of microcrystallinity in silicon semiconductor devices. *Appl. Spectrosc.*, 55(3):257–266, 2001.

[140] CM Hodges and J Akhavan. The use of Fourier-transform Raman-spectroscopy in the forensic identification of illicit drugs and explosives. *Spectrochim Acta A*, 46(2):303–307, 1990.

[141] EA Cutmore and PW Skett. Application of Fourier-transform Raman-spectroscopy to a range of compounds of pharmaceutical interest. *Spectrochim Acta A*, 49(5-6):809–818, 1993.

[142] Y Ozaki, R Cho, K Ikegaya, S Muraishi, and K Kawauchi. Potential of near-infrared Fourier-transform Raman-spectroscopy in food analysis. *Appl. Spectrosc.*, 46(10):1503–1507, 1992.

[143] TG Spiro (Ed.). Biological Applications of Raman Spectroscopy, Vol. 3. *John Wiley, New York, NY*, 1988.

[144] HU Gremlich and B Yan (Eds.). Infrared and Raman Spectroscopy of Biological Materials. *Marcel Dekker, Inc., New York, NY*, 2001.

[145] B Barry and R Mathies. Resonance Raman microscopy of rod and cone photoreceptors. *J. Cell. Biol.*, 94(2):479–482, 1982.

[146] M Gniadecka, OF Nielsen, DH Christensen, and HC Wulf. Structure of water, proteins, and lipids in intact human skin, hair, and nail. *J. Invest. Dermatol.*, 110(4):393–398, 1998.

[147] MT Kirchner, HGM Edwards, D Lucy, and AM Pollard. Ancient and modern specimens of human teeth: A Fourier transform Raman spectroscopic study. *J. Raman Spectrosc.*, 28(2-3):171–178, 1997.

[148] VE de Oliveira, HV Castro, HGM Edwards, and LFC de Oliveira. Carotenes and carotenoids in natural biological samples: a Raman spectroscopic analysis. *J. Raman Spectrosc.*, 41(6):642–650, 2009.

[149] M Schulmerich, J Cole, K Dooley, M Morris, J Kreider, S Goldstein, S Srinivasan, and B Pogue. Noninvasive Raman tomographic imaging of canine bone tissue. *J. Biomed. Opt.*, 13(2):020506, 2008.

[150] DA Troyer, Y Tang, R Bedolla, SG Adhvaryu, IM Thompson, S Abboud-Werner, L-Z Sun, WE Friedrichs, and LA deGraffenried. Characterization of PacMetUT1, a recently isolated human prostate cancer cell line. *Prostate*, 68(8):883–92, 2008.

[151] EM Goldblatt, ER Gentry, MJ Fox, SM Gryaznov, C Shen, and BS Herbert. The telomerase template antagonist GRN163L alters MDA-MB-231 breast cancer cell morphology, inhibits growth, and augments the effects of paclitaxel. *Mol. Cancer Ther.*, 8(7):2027–35, 2009.

[152] M Igawa, T Tanabe, GW Chodak, and DB Rukstalis. N-(4-hydroxyphenyl) retinamide induces cell cycle specific growth inhibition in PC3 cells. *Prostate*, 24(6):299–305, 1994.

[153] KR Stone, DD Mickey, H Wunderli, GH Mickey, and DF Paulson. Isolation of a human prostate carcinoma cell line (DU 145). *Int. J. Cancer*, 21(3):274–81, 1978.

[154] A Tomida, J Yun, and T Tsuruo. Glucose-regulated stresses induce resistance to camptothecin in human cancer cells. *Int. J. Cancer*, 68(3):391–396, 1996.

[155] YK Zhang, N Fujita, and T Tsuruo. Caspase-mediated cleavage of p21(Wafl)/(Cip1) converts cancer cells from growth arrest to undergoing apoptosis. *Oncogene*, 18:1131–1138, 1999.

[156] A Compaan, MC Lee, and GJ Trott. Phonon populations by nanosecond-pulsed Raman-scattering in Si. *Phys. Rev. B.*, 32:6731–6741, 1985.

[157] A Jirasek, Q Matthews, M Hilts, G Schulze, MW Blades, and RFB Turner. Investigation of a 2D two-point maximum entropy regularization method for signal-to-noise ratio enhancement: application to CT polymer gel dosimetry. *Phys. Med. Biol.*, 51:2599–2617, 2006.

[158] Q Matthews, A Jirasek, N Virji-Babul, A Babul, and T Cheung. Investigation of a two-point maximum entropy regularization method for signal enhancement applied to magnetoencephalography data. *Biomed. Signal Process. Contr.*, 3:78–87, 2008.

[159] LS Greek, HG Schulze, MW Blades, AV Bree, BB Gorzalka, and RFB Turner. SNR enhancement and deconvolution of Raman-spectra using a 2-point entropy regularization method. *Appl. Spectrosc.*, 49(4):425–431, 1995.

[160] HG Schulze, RB Foist, AI Jirasek, A Ivanov, and RFB Turner. Two-point maximum entropy noise discrimination in spectra over a range of baseline offsets and signal-to-noise ratios. *Appl. Spectrosc.*, 61:157–164, 2007.

[161] CE Shannon. A mathematical theory of communication. *Bell Syst. Tech. J.*, 27(3):379–423, 1948.

[162] A Savitzky and MJE Golay. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.*, 36:1627, 1964.

[163] IT Jolliffe. Principal Component Analysis. *Springer-Verlag New York, Inc., New York, NY*, 2002.

[164] JS Milton and JO Tsokos. Statistical methods in the biological and health sciences. *McGraw Hill, New York, NY*, 1983.

[165] CA Lieber and A Mahadevan-Jansen. Automated method for subtraction of fluorescence from biological Raman spectra. *Appl. Spectrosc.*, 57(11):1363–7, 2003.

[166] CM Krishna, GD Sockalingum, RA Bhat, L Venteo, P Kushtagi, M Pluot, and M Manfait. FTIR and Raman microspectroscopy of normal, benign, and

malignant formalin-fixed ovarian tissues. *Anal. Bioanal. Chem.*, 387(5):1649–1656, 2007.

[167] H Wang, N Huang, J Zhao, H Lui, M Korbelik, and H Zeng. Depth-resolved in vivo micro-Raman spectroscopy of a murine skin tumor model reveals cancer-specific spectral biomarkers. *J. Raman Spectrosc.*, 42(2):160–166, 2011.

[168] Q Matthews, A Jirasek, J Lum, X Duan, and AG Brolo. Variability in Raman spectra of single human tumor cells cultured in vitro: correlation with cell cycle and culture confluency. *Appl. Spectrosc.*, 64(8):871–87, 2010.

[169] JR Mourant, RR Gibson, TM Johnson, S Carpenter, KW Short, YR Yamada, and JP Freyer. Methods for measuring the infrared spectra of biological cells. *Phys. Med. Biol.*, 48(2):243–257, 2003.

[170] S Koljenović, TC Bakker Schut, R Wolthuis, B de Jong, L Santos, PJ Caspers, JM Kros, and GJ Puppels. Tissue characterization using high wave number Raman spectroscopy. *J. Biomed. Opt.*, 10(3):031116, 2005.

[171] ML Whitfield, LX Zheng, A Baldwin, T Ohta, MM Hurt, and WF Marzluff. Stem-loop binding protein, the protein that binds the $3'$ end of histone mRNA, is cell cycle regulated by both translational and posttranslational mechanisms. *Mol. Cell Biol.*, 20(12):4188–4198, 2000.

[172] ML Whitfield, G Sherlock, AJ Saldanha, JI Murray, CA Ball, KE Alexander, JC Matese, CM Perou, MM Hurt, PO Brown, and D Botstein. Identification of genes periodically expressed in the human cell cycle and their expression in tumors. *Mol. Biol. Cell*, 13(6):1977–2000, 2002.

[173] DM Carey and GM Korenowski. Measurement of the Raman spectrum of liquid water. *J. Chem. Phys.*, 108(7):2669–2675, 1998.

[174] M Levitt. Conformational preferences of amino-acids in globular proteins. *Biochem.*, 17(20):4277–4284, 1978.

[175] S Cooper. Rethinking synchronization of mammalian cells for cell cycle analysis. *Cell Mol. Life Sci.*, 60(6):1099–1106, 2003.

[176] S Cooper, G Iyer, M Tarquini, and P Bissett. Nocodazole does not synchronize cells: implications for cell-cycle control and whole-culture synchronization. *Cell Tissue Res.*, 324(2):237–42, 2006.

[177] S Boydston-White, T Gopen, S Houser, J Bargonetti, and M Diem. Infrared spectroscopy of human tissue. V. Infrared spectroscopic studies of myeloid leukemia (ML-1) cells at different phases of the cell cycle. *Biospectroscopy*, 5(4):219–27, 1999.

[178] Q Matthews, AG Brolo, J Lum, X Duan, and A Jirasek. Raman spectroscopy of single human tumour cells exposed to ionizing radiation in vitro. *Phys. Med. Biol.*, 56(1):19–38, 2011.

[179] V Vucic, ER Isenovic, M Adzic, S Ruzdijic, and MB Radojcic. Effects of gamma-radiation on cell growth, cycle arrest, death, and superoxide dismutase expression by DU145 human prostate cancer cells. *Braz. J. Med. Biol. Res.*, 39(2):227–236, 2006.

[180] N Gault, O Rigaud, JL Poncy, and JL Lefaix. Infrared microspectroscopy study of γ-irradiated and H_2O_2-treated human cells. *Int. J. Radiat. Biol.*, 81(10):767–79, 2005.

[181] N Gault, O Rigaud, JL Poncy, and JL Lefaix. Biochemical alterations in human cells irradiated with α-particles delivered by macro- or microbeams. *Radiat. Res.*, 167(5):551–562, 2007.

[182] AD Meade, C Clarke, HJ Byrne, and FM Lyng. Fourier Transform Infrared Microspectroscopy and Multivariate Methods for Radiobiological Dosimetry. *Radiat. Res.*, 173(2):225–237, 2010.

[183] GP Amorino, ML Freeman, and H Choy. Enhancement of radiation effects in vitro by the estrogen metabolite 2-methoxyestradiol. *Radiat. Res.*, 153(4):384–91, 2000.

[184] SY Park, YM Kim, and H Pyo. Gefitinib radiosensitizes non-small cell lung cancer cells through inhibition of ataxia telangiectasia mutated. *Mol. Cancer*, 9:222, 2010.

[185] A Sak, M Stuschke, R Wurm, G Schroeder, B Sinn, G Wolf, and V Budach. Selective inactivation of DNA-dependent protein kinase with antisense oligodeoxynucleotides: consequences for the rejoining of radiation-induced DNA double-strand breaks and radiosensitivity of human cancer cell lines. *Cancer Res.*, 62(22):6621–4, 2002.

[186] BW Robinson and DS Shewach. Radiosensitization by gemcitabine in p53 wild-type and mutant MCF-7 breast carcinoma cell lines. *Clin. Cancer Res.*, 7(8):2581–9, 2001.

[187] RK Schmidt-Ullrich, K Valerie, W Chan, DE Wazer, and PS Lin. Expression of oestrogen receptor and transforming growth factor-alpha in MCF-7 cells after exposure to fractionated irradiation. *Int. J. Radiat. Biol.*, 61(3):405–15, 1992.

[188] A Hahnel, H Wichmann, M Kappler, M Kotzsch, D Vordermark, H Taubert, and M Bache. Effects of osteopontin inhibition on radiosensitivity of MDA-MB-231 breast cancer cells. *Radiat. Oncol.*, 5:82, 2010.

[189] A Wouters, B Pauwels, HAJ Lambrechts, GGO Pattyn, J Ides, M Baay, P Meijnders, F Lardon, and JB Vermorken. Counting clonogenic assays from normoxic and anoxic irradiation experiments manually or by using densitometric software. *Phys. Med. Biol.*, 55(7):N167–78, 2010.

[190] TL DeWeese, JM Shipman, LE Dillehay, and WG Nelson. Sensitivity of human prostatic carcinoma cell lines to low dose rate radiation exposure. *J. Urol.*, 159(2):591–8, 1998.

[191] PJ Colletier, F Ashoori, D Cowen, RE Meyn, P Tofilon, ME Meistrich, and A Pollack. Adenoviral-mediated p53 transgene expression sensitizes both wild-type and null p53 prostate cancer cells in vitro to radiation. *Int. J. Radiat. Oncol. Biol. Phys.*, 48(5):1507–12, 2000.

[192] D Chendil, RS Ranga, D Meigooni, S Sathishkumar, and MM Ahmed. Curcumin confers radiosensitizing effect in prostate cancer cell line PC-3. *Oncogene*, 23(8):1599–607, 2004.

[193] JZ Wang, JG Rhee, P Shi, RD Stewart, and X Allen Li. In vitro determination of radiation sensitivity parameters for DU-145 prostate cancer cells. *Int. J. Radiat. Biol.*, 84(6):515–22, 2008.

[194] NE Fullerton, M Boyd, RJ Mairs, WN Keith, O Alderwish, MM Brown, A Livingstone, and D Kirk. Combining a targeted radiotherapy and gene therapy approach for adenocarcinoma of prostate. *Prostate Cancer Prostatic Dis.*, 7(4):355–63, 2004.

[195] CJ Rosser, M Tanaka, LL Pisters, N Tanaka, LB Levy, DC Hoover, HB Grossman, TJ Mcdonnell, DA Kuban, and RE Meyn. Adenoviral-mediated PTEN transgene expression sensitizes Bcl-2-expressing prostate cancer cells to radiation. *Cancer Gene Ther.*, 11(4):273–9, 2004.

[196] JB Mitchell, R Choudhuri, K Fabre, AL Sowers, D Citrin, SD Zabludoff, and JA Cook. In vitro and in vivo radiation sensitization of human tumor cells by a novel checkpoint kinase inhibitor, AZD7762. *Clin. Cancer Res.*, 16(7):2076–84, 2010.

[197] L Hui, Y Zheng, Y Yan, J Bargonetti, and DA Foster. Mutant p53 in MDA-MB-231 breast cancer cells is stabilized by elevated phospholipase D activity and contributes to survival signals generated by phospholipase D. *Oncogene*, 25(55):7305–10, 2006.

[198] JR Williams, Y Zhang, H Zhou, DS Gridley, CJ Koch, JM Slater, and JB Little. Overview of radiosensitivity of human tumor cells to low-dose-rate irradiation. *Int. J. Radiat. Oncol. Biol. Phys.*, 72(3):909–17, 2008.

[199] Wulf Dröge. Free radicals in the physiological control of cell function. *Physiol. Rev.*, 82(1):47–95, 2002.

[200] H Hermeking, C Lengauer, K Polyak, TC He, L Zhang, S Thiagalingam, KW Kinzler, and B Vogelstein. 14-3-3 sigma is a p53-regulated inhibitor of G2/M progression. *Mol. Cell.*, 1(1):3–11, 1997.

[201] B Xiao, SJ Smerdon, DH Jones, GG Dodson, Y Soneji, A Aitken, and SJ Gamblin. Structure of a 14-3-3 protein and implications for coordination of multiple signalling pathways. *Nature*, 376(6536):188–91, 1995.

[202] MJ van Hemert, HY Steensma, and GP van Heusden. 14-3-3 proteins: key regulators of cell division, signalling and apoptosis. *BioEssays*, 23(10):936–46, 2001.

[203] L Chantalat, DA Skoufias, JP Kleman, B Jung, O Dideberg, and RL Margolis. Crystal structure of human survivin reveals a bow tie-shaped dimer with two unusual alpha-helical extensions. *Mol. Cell*, 6(1):183–9, 2000.

[204] A Chakravarti, GG Zhai, M Zhang, R Malhotra, DE Latham, MA Delaney, P Robe, U Nestler, Q Song, and J Loeffler. Survivin enhances radiation resistance in primary human glioblastoma cells via caspase-independent mechanisms. *Oncogene*, 23(45):7494–506, 2004.

[205] K Kami, R Doi, M Koizumi, E Toyoda, T Mori, D Ito, Y Kawaguchi, K Fujimoto, M Wada, S-I Miyatake, and M Imamura. Downregulation of survivin by siRNA diminishes radioresistance of pancreatic cancer cells. *Surgery*, 138(2):299–305, 2005.

[206] F Rödel, J Hoffmann, L Distel, M Herrmann, T Noisternig, T Papadopoulos, R Sauer, and C Rödel. Survivin as a radioresistance factor, and prognostic and therapeutic target for radiotherapy in rectal cancer. *Cancer Res.*, 65(11):4881–7, 2005.

[207] Z Khan, N Khan, RP Tiwari, IK Patro, GBKS Prasad, and PS Bisen. Downregulation of survivin by oxaliplatin diminishes radioresistance of head and neck squamous carcinoma cells. *Radiother. Oncol.*, 96(2):267–73, Aug 2010.

[208] MA Chaudhry, B Kreger, and RA Omaruddin. Transcriptional modulation of micro-RNA in human cells differing in radiation sensitivity. *Int. J. Radiat. Biol.*, 86(7):569–83, 2010.

[209] BA Foster, JR Gingrich, ED Kwon, C Madias, and NM Greenberg. Characterization of prostatic epithelial cell lines derived from transgenic adenocarcinoma of the mouse prostate (TRAMP) model. *Cancer Res.*, 57(16):3325–30, 1997.

[210] TJ Jorgensen, S Katz, EK Wittmack, S Varghese, T Todo, SD Rabkin, and RL Martuza. Ionizing radiation does not alter the antitumor activity of herpes simplex virus vector G207 in subcutaneous tumor models of human and murine prostate cancer. *Neoplasia*, 3(5):451–6, 2001.

[211] CJ Frank, RL McCreery, and DC Redd. Raman spectroscopy of normal and diseased human breast tissues. *Anal. Chem.*, 67(5):777–83, 1995.

[212] H Shafiee, MB Sano, EA Henslee, JL Caldwell, and RV Davalos. Selective isolation of live/dead cells using contactless dielectrophoresis (cDEP). *Lab Chip*, 10(4):438–45, 2010.

[213] J Kneipp, H Kneipp, M McLaughlin, D Brown, and K Kneipp. In vivo molecular probing of cellular compartments with gold nanoparticles and nanoaggregates. *Nano Lett.*, 6(10):2225–2231, 2006.

[214] A Shamsaie, M Jonczyk, J Sturgis, JP Robinson, and J Irudayaraj. Intracellularly grown gold nanoparticles as potential surface-enhanced Raman scattering probes. *J. Biomed. Opt.*, 12(2):020502, 2007.

[215] KL Nowak-Lovato and KD Rector. Targeted surface-enhanced Raman scattering nanosensors for whole-cell pH imagery. *Appl. Spectrosc.*, 63(4):387–395, 2009.

[216] CL Zavaleta, BR Smith, I Walton, W Doering, G Davis, B Shojaei, MJ Natan, and SS Gambhir. Multiplexed imaging of surface enhanced Raman scattering nanotags in living mice using noninvasive Raman spectroscopy. *P. Natl. Acad. Sci. USA*, 106(32):13511–13516, 2009.

[217] CT Nguyen, JT Nguyen, S Rutledge, J Zhang, C Wang, and GC Walker. Detection of chronic lymphocytic leukemia cell surface markers using surface enhanced Raman scattering gold nanoparticles. *Cancer Lett.*, 292(1):91–7, 2010.

Subject: reproduction of published material in my doctoral thesis
From: Quinn Matthews <qmatthew@uvic.ca>
Date: Fri, 3 Jun 2011 14:48:43 -0700
To: Office@s-a-s.org

Dear SAS,
I am currently preparing my doctoral thesis, tentatively entitled "Radiobiological analysis of tumour cells with single-cell Raman spectroscopy", to be published by the University of Victoria, British Columbia, Canada.

I would appreciate permission to reproduce, in my thesis, the majority of the text and figures from the article "Variability in Raman spectra of single human tumor cells cultured in vitro: correlation with cell cycle and culture confluency", which is published in the SAS journal Applied Spectroscopy, volume 64(8), pp. 871-887, 2010. I am the lead author of the article, and I have permission from the co-authors to use the material.

Thank you for your prompt attention to this request. Please inform me if you require any other information from myself or my co-authors with regard to this request.

Sincerely,
Quinn Matthews

Ph.D. Candidate,
Department of Physics and Astronomy
University of Victoria
qmatthew@uvic.ca

Permission granted for the use requested above:
Permission granted this 16th **day of**
June **2011. Full citation required.**
[redacted]
Bonnie Saylor, Executive Director, SAS

To: permissions@iop.org,
Cc:
Bcc:
Subject: reproduction of published material in my doctoral thesis
From: Quinn Matthews <qmatthew@uvic.ca> - Friday 03/06/2011 22:41
Sent by: quinnmatthews@gmail.com

Dear IOP Publishing,

I am currently preparing my doctoral thesis, tentatively entitled "Radiobiological analysis of tumour cells with single-cell Raman spectroscopy", to be published by the University of Victoria, British Columbia, Canada.

I would appreciate permission to reproduce the majority of the text and figures from the article "Raman spectroscopy of single human tumour cells exposed to ionizing radiation in vitro", which is published in the IOP journal Physics in Medicine and Biology, volume 56, pp. 19-38, 2011. I am the lead author of the article, and I have permission from the co-authors to use the material.

Thank you for your prompt attention to this request. Please inform me if you require any other information from myself or my co-authors with regard to this matter.

Sincerely,
Quinn Matthews

Ph.D. Candidate,
Department of Physics and Astronomy
University of Victoria
qmatthew@uvic.ca

PERMISSION TO REPRODUCE AS REQUESTED IS GIVEN PROVIDED THAT:

~~(a) the consent of the author(s) is obtained~~
(b) the source of the material including author, title of article, title of journal, volume number, issue number (if relevant), page range (or first page if this is the only information available), date and publisher is acknowledged.
(c) for material being published electronically, a link back to the original article should be provided (via DOI).

IOP Publishing Ltd
Dirac House
Temple Back
BRISTOL
BS1 6BE

6/6/2011
Date

[redacted]
Rights & Permissions