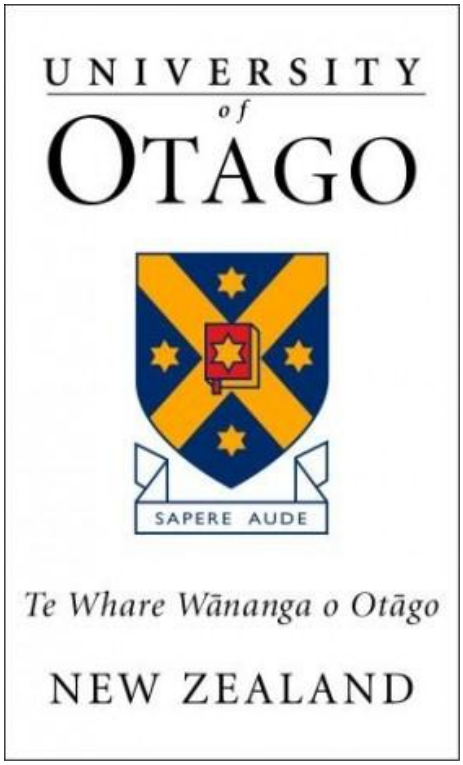
# A Raman Spectroscopic Study of Paint and Dairy Samples

Geoffrey Smith




A thesis submitted for the degree of

Master of Science

At the University of Otago, Dunedin,

2012

# Abstract

This thesis investigates the use of Raman spectroscopic techniques for analysing paint and dairy samples (milk powder and model cheese systems).

Several samples of paint from a painting called ‘Hope’ by Edward Burne-Jones were obtained from the Dunedin Public Art Gallery. The purpose of this study was to determine whether the pigments and/or binder could be identified using Raman microscopy. The technique conclusively identified several pigments in two of the samples including: cerulean blue, Prussian blue, Vermilion and zinc white.

FT-Raman spectroscopic data was investigated with respect to its ability to be used to identify milk powder samples spiked with calcium carbonate and whether quantification of the adulterant is possible. Use of the spectroscopic data with principal components analysis (PCA) was useful for identifying samples which had been spiked and for distinguishing samples with larger differences in calcium carbonate concentration. Partial least squares (PLS) and univariate analysis were able to be used with spectroscopic data to create standard curve plots which could be used in the quantification of calcium carbonate in milk powder. However, limit of quantification (LOQ) values indicated that neither technique would be effective in quantifying calcite concentrations as low as those used for calcium fortification in retail milk powder samples.

Model cheese systems were the primary focus of this research. Raman spectroscopic techniques (primarily Raman microscopy) were investigated with respect to their uses in studying cheese. FT-Raman spectroscopy with PCA was used determine whether model bi-component cheese mixtures could be distinguished based on constituent concentrations. PCA was found to be very effective for achieving this. Mixtures were then analysed with respect to whether interactions between components were observable in the spectra (as band broadening or shifting). Interactions were not found. Morphology of the model cheese mixtures was investigated using confocal Raman microscopy. This technique was effective for analysis of morphology of model bi-component cheese mixtures in both two and three dimensions.

Overlap between peaks of different components is the main limitation of this method when analysing morphology. Raman microscopy was compared to confocal laser scanning microscopy (CLSM) with respect to production of 3D images. While CLSM produced images in much less time, Raman microscopy has the advantage of being able to identify carbohydrates.

# Acknowledgements

First and foremost, I would like to thank my supervisor, Keith Gordon, for all his support and for making two years of hard study and research as enjoyable as it possibly could be. I would also like to thank my co-supervisor Allan Blackman for staying composed when I ambushed him with decent chunks of thesis to critique. My thanks goes to Steve Holroyd for providing the samples, all his support on my ventures to Palmerston North and for offering feedback on several sections of this thesis. I would like to say a big thank you to Liz Nickless for providing samples, for spending time showing me the analytical techniques employed by the dairy industry and for all the help she gave whenever I enquired. I would also like to thank her for all the images she provided for use in this thesis and for letting me use her image processing software. Thank you to Andrew McNaughton for taking the time to help me collect images using the confocal laser scanning microscope in Dunedin. My thanks go to Austina Clark for her help with all my statistics related issues and for performing the analysis of variance on the milk powder data.

I would like to thank the KCG Empire (Raphael, Matt, Jacqui, Sara, Mike, Sam, Stasi, Holly, Adamina and Alvie). I could not have asked for a more pleasant group to be involved with. I owe Jacqui thanks for all her help when the microscope would go on strike. Thank you, Sara, for your patience on each occasion where I commandeered your computer to use The Unscrambler.

I would also like to thank Sijing Feng for his help in the proof reading process. My penultimate thanks go to Carol and Trevor for all the support they have given while I have been at University and all the years prior. It has made the entire university experience a very enjoyable one.

Finally, I would like to thank the Otago University Chemistry Department. I could not think of a better department to have been part of.

# Table of Contents

Abstract……i
Acknowledgements……...….iii
Table of Contents……v
Expanded Table of Contents…….vii
List of Abbreviations……..xi
Chapter 1 Introduction…….……1
Chapter 2 Experimental……...……..21
Chapter 3 Paint……..……33
Chapter 4 Milk Powder…….……45
Chapter 5 Cheese Analysis……..…….…73
Chapter 6 Raman Mapping of Model Cheese Systems…….……83
Chapter 7 Future Work……107
References……...……109
Appendix A Stage Movement in the Z-Axis……...113
Appendix B Contour Maps from 3D Mapping……..……..115

# Expanded Table of Contents


Abstract……...……i

Acknowledgements……..……....……..iii

Table of Contents…….....….……..v

Expanded Table of Contents…….....…….vii

List of Abbreviations…….....……xi

Chapter 1 Introduction…….....……..….1

1.1 Raman Spectroscopy…….....……....1

1.1.1 Raman microscopy…….....…….....2

1.2 Confocal Laser Scanning Microscopy……..…….....…….4

1.3 Paint Samples…….....….....……..4

1.4 Dairy Samples……...….....…..11

1.5 Thesis Aims……...….....……..18

Chapter 2 Experimental.....21

2.1 Instrumentation.....21

2.1.1 Raman microscope.....21

2.1.2 FT-Raman Spectrometer.....21

2.1.3 Confocal Laser Scanning Microscope (CLSM).....21

2.2 Software.....21

2.3 Paint, Skin and Oil Sample Analysis.....22

2.3.1 Paint fleck 1 analysis using the Raman microscope.....22

2.3.2 Paint fleck 2 analysis using the Raman microscope.....23

2.3.3 Paint fleck 3 analysis using the Raman microscope.....23

2.3.4 Skin fleck analysis using the Raman microscope.....23

2.3.5 Linseed oil analysis using the FT-Raman spectrometer.....23

2.4 Analysis of Milk Powder Samples.....23

2.5Cheese Sample Preparation.....24

2.5.1 Preparation of binary mixture set A (BMA).....24

2.5.2 Preparation of binary mixture set B (BMB).....25
2.5.3 Preparation of tertiary cheese component mixtures.....25
2.6 Cheese Analysis.....25
2.6.1 Distinguishing BMA mixtures using FT-Raman spectroscopy.....25
2.6.2 Identification of interactions within 1:1 binary mixtures.....26
2.6.3 Tertiary mixture analysis.....26
2.7 Producing Contour Maps.....26
2.7.1 Producing the 2D contour maps.....26
2.7.2 Producing the 3D contour maps.....30
2.8 Producing images using CLSM.....31
Chapter 3 Paint.....33
3.1 Introduction.....33
3.2 Identification of Pigments.....33
3.2.1 Paint fleck 1.....34
3.2.2 Paint fleck 2.....38
3.2.3 Paint fleck 3.....41
3.3 Identification of Binder.....42
3.4 Conclusions.....43
Chapter 4 Milk Powder.....45
4.1 Introduction.....45
4.2 Standard Normal Variate (SNV) Pre-Processing.....46
4.3 Principal Component Analysis (PCA).....46
4.3.1 Introduction to the technique.....46
4.3.2 Results and Discussion.....47
4.4 Univariate Analsysis.....55
4.4.1 Introduction to the technique.....55
4.4.2 Results and discussion.....55
4.4.3 Analysis of variance.....57
4.5 Partial Least Squares (PLS) Analysis.....63
4.5.1 Introduction to the technique.....63
4.5.2 Results and discussion.....66

4.5.3 Comparison of PLS with univariate analysis.....70
4.6 Conclusions.....70
Chapter 5 Cheese Component Analysis.....73
5.1 Introduction .....73
5.2 Differentiation of Cheese Mixtures.....75
5.3 Interactions within 1:1 Binary Mixtures.....79
5.4 Interactions within Tertiary Mixtures.....80
5.5 Conclusions.....82
Chapter 6 Raman Mapping of Model Cheese Systems.....83
6.1 Introduction.....83
6.2 Results and Discussion of 2D Contour Maps.....85
6.3 Results and Discussion of 3D Contour Maps.....89
6.4 Alternative Standardisation Method.....97
6.5 High Resolution Contour Map.....99
6.6 Confocal Laser Scanning Microscopy (CLSM) Results and Discussion.....101
6.7 CLSM Versus Raman Microscopy.....103
6.8 Conclusions.....105
Chapter 7 Future Work.....107
References.....109
Appendix A Stage Movement in the Z-Axis.....113
Appendix B Contour Maps from 3D Mapping.....115

# List of Abbreviations

| | |
|---|---|
| 2D: | 2 dimensional |
| 3D: | 3 dimensional |
| AMF: | anhydrous milk fat |
| AN: | area normalisation |
| ANOVA: | analysis of variance |
| BL: | baseline |
| BMA: | binary mixture set A |
| BMB: | binary mixture set B |
| CC: | calcium carbonate |
| CCD: | charge coupled device |
| CD: | *c*-diene |
| CI: | confidence interval |
| CLSM: | confocal laser scanning microscopy |
| CS: | cornstarch |
| CTD: | *c*, *t*-diene |
| DF: | degrees of freedom |
| ELISA: | enzyme-linked immunosorbent assay |
| FG: | fat globule |
| FTIR: | Fourier-transform infrared |
| FT-Raman: | Fourier-transform Raman |
| HLS: | hard long saturates |
| HSS: | hard short saturates |
| HPLC: | high performance liquid chromatography |
| JK: | jack-knifing |
| LM: | long monounsaturated |
| LOD: | limit of detection |
| LOQ: | limit of quantification |
| MFG: | milk fat globule |

| | |
|---|---|
| NA: | numerical aperture |
| NIR: | near infrared |
| NMR: | nuclear magnetic resonance |
| Origin: | OriginPro (version 7.5, OriginLab Corporation, Massachusetts) |
| PC: | principal component |
| PCA: | principal component analysis |
| PLS: | partial least squares |
| RC: | rennet casein |
| RM: | Raman microscope/microscopy |
| RMSEC: | root mean square error of calibration |
| RMSECV: | root mean square error of cross validation |
| RMSEP: | root mean square error of prediction |
| SFC: | solid fat content |
| SM: | short monounsaturated |
| SNV: | standard normal variate |
| SS: | sample set |
| SSat: | short saturated |
| SSq: | sum of squares |
| SUC: | sucrose |
| TG: | triacylglycerol |
| TSC: | trisodium citrate |
| TM: | *t*-monounsaturated |

# Chapter 1 Introduction

The identification of unknown substances is an important part of analytical chemistry. There is a variety of techniques can be used for identification purposes. Four examples are nuclear magnetic resonance (NMR), infrared (IR) spectroscopy, Raman spectroscopy, and ultraviolet-visible (UV-vis) spectroscopy.

Paint and dairy products are two areas in which identification of constituents is important. Paint samples are of interest because identification of pigments can give information on the authenticity of an artwork.[1] Distinguishing between pigments can also have applications in forensic investigations.[2] Needless to say, knowing the content of food is also important. Analytical techniques are commonly employed in the monitoring of food production. Cheese in particular, has previously been studied with respect to whether near infra-red (NIR) spectroscopy and Raman spectroscopy could be useful in establishing its content, determining its structure and monitoring the manufacturing process.[3, 4, 5] This thesis investigates the applicability of Raman spectroscopy in both paint and dairy samples.

## 1.1 Raman Spectroscopy

The principle behind Raman techniques is called the Raman effect, which involves the inelastic scattering of laser photons by molecules. The energy loss of the scattered photons corresponds to the vibrational energy levels of the interacting molecule. Stokes, anti-Stokes or Raleigh scattering can be observed when a photon interacts with a molecule depending on whether the molecule relaxes to an energy level higher than, lower than, or the same as the initial energy state, respectively (**figure 1.1**). This is seen on the spectrum as an emitted photon whose frequency is either lower than, greater than or the same as that of the incident photon. The Raman effect is inherently weak with features such as stokes scattering having an intensity of $10^{-6}$ relative to that of the incident excitation.[6] The intensity of the bands is determined by the change in polarizability, of the molecule as it vibrates. Just as IR spectroscopy requires a change in dipole moment, if there is no change in the polarizabilty, there will be no observable bands.[6, 7] Raman spectroscopy can therefore give information on

the bonds present within a molecule. Utilising this information can allow substances such as pigments within a painting to be identified.

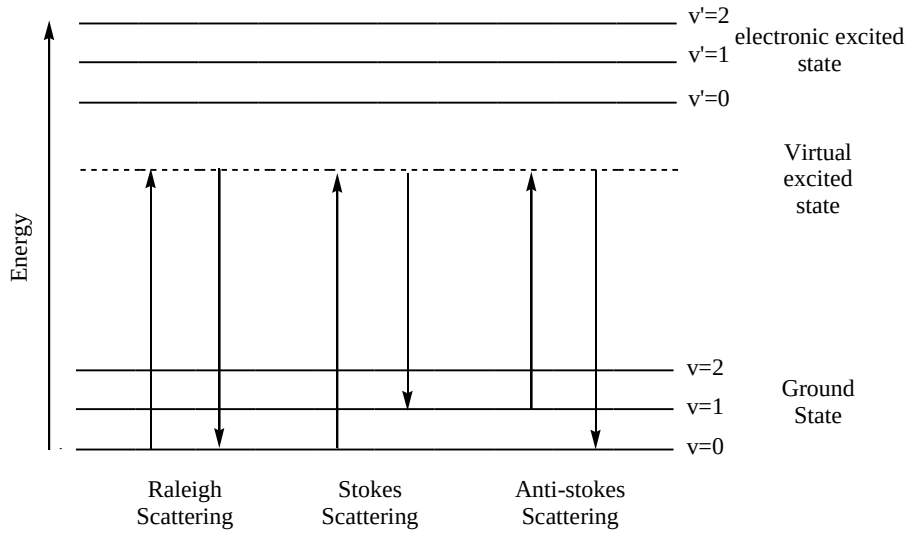



**Figure 1.1 Diagram illustrating Raman scattering which occurs when the sample is excited by an incident laser.**

There are two issues which can make collecting Raman spectra problematic. The first of these is fluorescence, a broad spectral emission which can inundate the Raman signal. The second is the weakness of the Raman effect. The weak Raman signals mean that noise, which can come from a variety of sources, can make the acquisition of good quality data difficult.

Fourier transform Raman (FT-Raman) spectroscopy offers improvements relating to the analytical problems of conventional Raman spectroscopy. FT-Raman spectroscopy utilises a Fourier transform operation to generate Raman spectra from raw data collected by the spectrometer. This imparts the advantage of high wavenumber precision, thus providing an improved ability to perform high resolution experiments.[7] Furthermore, the technique uses a 1064 nm laser wavelength. This near infrared wavelength allows spectra to be obtained without interference from fluorescence.[7]

### 1.1.1 Raman microscopy

Raman microscopy involves feeding a laser through a microscope objective onto a sample, resulting in Raman scattering. Backscattered light then passes through the microscope

objective where it is then directed to a charge-coupled device (CCD) detector. This technique allows Raman spectra to be collected from structures on a micrometer scale. A Senterra confocal Raman microscope is pictured in **figure 1.2**. Raman microscopy can be used for mapping samples to allow distribution of constituents within a sample to be observed. Mapping differs from imaging primarily in the method in which it is achieved. Mapping requires spectra to be taken at several points in a grid formation and the spectra are analysed with respect to the constituents present at each point. Imaging is performed by collecting several spectra simultaneously using a two dimensional detector.

Applications of Raman microscopy are numerous. Artworks and other valuable artefacts have been studied extensively using Raman microscopy in order to identify pigments and in some cases, determine authenticity. Some examples can be found in references 15-21 which are discussed in **section 1.3**. Medicines are also analysed using this technique. Mapping using Raman microscopy has been used to study drug distribution within the solid lipid matrices and drug dissolution behaviour of pharmaceuticals.[8] This was performed with the intent to provide a better understanding of the drug release behaviour. Results indicated that the technique did not have sufficient spatial resolution to determine the exact component distribution at the surface of extrudates; however, the technique was useful for chemically resolved imaging of drug distribution after dissolution. Mapping using Raman microscopy has also been used to study the distribution of the constituents within cheese samples.[9] Two methods for mapping constituents were used which utilised either univariate or multivariate analysis. Results showed that univariate analysis was effective at identifying component distribution in highly heterogeneous samples but was less effective in homogeneous samples. Multivariate analysis was effective for both homogeneous and heterogeneous samples. This study is described in more detail in **section 6.1**.

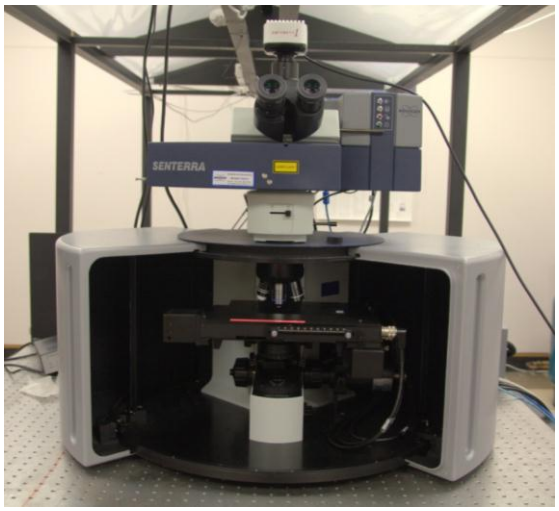



**Figure 1.2 Picture of a Senterra dispersive Raman microscope.**

## 1.2 Confocal Laser Scanning Microscopy (CLSM)

CLSM usually requires a sample to be stained with a fluorescent dye, provided the sample cannot fluoresce on its own and is able to be stained. Each component within the sample, whether it is protein, fat or other, requires a specific dye to stain it. Once stained, a laser is applied to the sample which excites the dye (not the sample itself) and causes fluorescence. The location of each component is identified by the location of the dyes within the sample. It is the confocal nature of the instrument which makes the technique useful for imaging. This is achieved through the aperture which has the purpose of filtering out signal from the areas outside the point of focus. This allows imaging of a sample at multiple depths due to the ability to obtain signal from well defined focal planes.[10] CLSM has vast numbers of applications in the observation and imaging of materials (outside the obvious biological applications). One such example involves observing the distribution of adhesives in paper.[11] CLSM has also been found to have applications in the imaging drug distribution within pharmaceutical systems.[12] Indeed, CLSM has already been applied for use on dairy systems, as discussed in **section 6.1**.

## 1.3 Paint Samples

There are three main components which generally make up most paints: binder, pigment and solvent. Each component has a different role in the painting process. Binder is used to bind pigments together, encourage adhesion of paint to a surface and ensure that when the paint

dries a continuous skin-like film is created. Pigment provides colour, opacity (protects the surface underneath) and helps the binder to stick. The solvent makes paint easier to apply and is usually quite volatile so it can evaporate, allowing the binder and pigment to stick to the surface. Solvent can be adjusted to change the drying rate to suit the method of paint application (for example, using a more volatile solvent in spray paint).[13] Varnish is then often applied to the dry paint in order to protect it from the environment. Thus, together with a number of binders which could be used, the chemical content of a painting can be highly variable. The date (and location) of discovery or development of binders and pigments is an invaluable 'fingerprint' which can be utilised to identify the origin of paintings. Clark has performed a large amount of work on paint samples using Raman microscopy and has investigated the ability of Raman microscopy to identify pigments and used his data to construe information about many pieces of artwork.[16-21] Therefore it is appropriate to refer to Clark's research while investigating samples of Burne-Jones' 'Hope' received from Dunedin Public Art Gallery.

When studying works of art, particularly those which involve paints, one must consider the amount of variation of chemicals which may be present. Four pigments, for example, which may be used in paints are zinc oxide (white pigment), copper aceto-arsenite (emerald green), $C_{20}H_{20}N_3Cl$ (magenta), and lead carbonate (white pigment).[14] This attests to the vast amount of chemical variability which may be found in samples, even when considering what appears to be the same pigment. Furthermore, the chemicals used to create certain colours has changed with time, which is a factor which is sometimes exploited for art authentication purposes.[15]

Possibly the most relevant of Clark's papers with respect to this thesis is the investigation into the pigments used by Portuguese artist Henrique Pousão in 23 of his paintings.[16] The goals of this investigation were to identify the pigments and to determine how varied the painter's palette was. This work involved the use of a He-Ne (632.8 nm) laser to provide Raman excitation.[16] Samples were obtained from paintings using a surgical scalpel. They were embedded in an epoxy polymeric resin, polished with silicon carbide and then with cloth using an alumina – water slurry. A total of 151 cross sections were collected from the 23 Pousão paintings. The cross sections were then subjected to examination by optical and Raman

microscopy. The samples were compared to reference samples collected from various locations. Raman microscopy led to the identification of 37 compounds each of which were placed into one of three categories: pigment, extender or trace component. A list of compounds found by Clark can be seen in **table 1.1**. The information gathered was found to be very effective for establishing Pousão's palette. This investigation illustrates the ability of Raman microscopy to be used for the analysis of art and identification of pigments.

It is important to note that Chapter 3 of this thesis shares this particular aim with the aforementioned paper. However, the goal is to identify the pigments used by Burne-Jones in his painting 'Hope' which is frequently displayed at the Dunedin Public Art Gallery. Similarly to the samples from Pousão's work, small flecks of paint were obtained using a surgical scalpel, however there was no prior preparation of the flecks of 'Hope'. Clark used Raman microscopy to elucidate the pigments used in other artefacts and discusses it in several other papers. The work includes the analysis of pigments on a 16$^{th}$ century book of poetry,[17] the identification of cinnabar on ancient Chinese manuscripts,[18] and the identification of pigments and gems on the Tours Gospel which were then compared to those present on other Gospels and bibles.[19]

**Table 1.1 A table which illustrates the range of pigments which have been identified in various paintings by Henrique Pousão using Raman microscopy.[16] The paintings are grouped into four chronological periods, Early, French, Italian and Final, according to the date or locality at which they were painted. The Early period refers to paintings executed in Portugal as a student of the Academia Portuense de Belas-Artes; the French and Italian periods refer to paintings executed in France and Italy, respectively; the Final period refers to paintings executed in the last year of his life on the return trip to Portugal.**

| | Painting | | | | | | | | | | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Early | | | French | | | Italian | | | | | | | | | | | | | | | Final | |
| **Pigments** | **A** | **B** | **C** | **D** | **E** | **F** | **G** | **H** | **I** | **J** | **K** | **L** | **M** | **N** | **O** | **P** | **Q** | **R** | **S** | **T** | **U** | **V** | **W** |
| ***White*** | | | | | | | | | | | | | | | | | | | | | | | |
| Anatase | • | | | | | | | | | | • | | • | | • | | | | | | | | |
| Barium white | | | | | | • | • | • | • | • | • | | • | • | | • | | • | | • | • | | |
| Basic lead sulfate | | | | | | | | | • | | • | | | | | | | • | | | | | |
| Calcium carbonate | | | • | | • | • | • | | • | | • | | | • | • | • | • | • | • | | | | |
| Kaolinite | | | | | | | | | | | | | | | | | | | | • | | | • |
| Lead carbonate | • | • | • | • | • | • | • | • | • | | • | | • | • | • | • | • | | | • | • | | • |
| Lead white | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • |
| Lead sulphate | | | • | | • | • | | | • | | • | • | | • | | | • | • | | • | | | • |
| Quartz | | | • | | | | • | | | | • | | | | | | | | | | | • | |
| Rutile | | | | | | | | | | | • | • | | | | | | | | | | | |
| Zinc white | | | | | | | | | | •[a] | | | | | | | | | | •[a] | | | • |
| ***Yellow*** | | | | | | | | | | | | | | | | | | | | | | | |
| Cadmium yellow | | | • | | | • | • | • | | | • | •[a] | • | | | | | | | | • | • | |
| Chrome yellow | • | | • | • | • | • | | • | • | | • | • | | • | • | • | • | • | • | • | • | | • |
| Iron(III) oxyhydroxide | | | | • | • | | • | | • | • | • | • | • | • | • | | | • | | | • | • | • |
| Lead antimonate yellow | | | | | • | • | | | • | | • | • | | • | | | | • | | • | | | |
| Strontium yellow | • | • | • | • | • | | | | | | | | | | | | | • | | | | | |
| Zinc yellow | | | | | | • | | • | | | • | • | • | | • | | | | | • | | • | |
| ***Orange*** | | | | | | | | | | | | | | | | | | | | | | | |
| Chrome orange | • | | • | • | • | • | • | | • | • | • | • | | | • | • | • | | • | | • | | |
| Realgar/parealgar | | | | • | | | • | • | | | | | | | | | | | | | | | • |
| ***Red*** | | | | | | | | | | | | | | | | | | | | | | | |
| Cadmium red | | | | | | | | | | • | | | | | | | | | | • | | | • |
| Cochineal lake | • | | • | • | • | • | | • | • | • | • | • | | • | • | | • | • | • | • | • | • | • |
| Iron(III) oxide | • | • | • | • | • | | • | • | • | • | • | • | • | • | • | | • | | • | • | • | • | • |
| Madder lake | | | | | | | | | | • | | | | | | | | | | | | | |
| Red lead | | | • | | | | | | | | | | | | | | | | | • | | | |
| Vermilion | • | | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • |
| ***Blue*** | | | | | | | | | | | | | | | | | | | | | | | |
| Cobalt blue | • | | • | | • | • | • | • | • | | • | | • | | | • | • | •[a] | • | • | | • | |
| Prussian blue | | | | • | • | • | | | • | • | • | • | | • | • | | | | | • | • | • | • |
| Ultramarine | • | | • | | • | • | • | • | • | • | • | • | | • | • | • | • | | | • | • | • | • |
| ***Green*** | | | | | | | | | | | | | | | | | | | | | | | |
| Brochantite | | | | | | | | • | | | | | | | | | | | | | | | |
| Celadonite | | | | | • | | | | • | | | | | | | | | | | | | | |
| Chrome green | | | | | | • | | • | | | | | | • | • | | | | | | • | | |
| Emerald green | | | | | | • | | • | | | • | • | | • | | | | | | | | • | |
| Malachite | | | | | | | | • | | | | | | | | | | | | | | | |
| Scheele's green | | | | | | | | • | | | • | | | | | | | | | | | | |
| Viridian | | | | • | • | • | | • | • | | | • | | | | | | | | • | | | • |
| ***Black*** | | | | | | | | | | | | | | | | | | | | | | | |
| Carbon-based black | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • | • |
| Copper sulphide | | | | | | | | | | | | | | • | | | | | | • | | | |

**A** – *Casa rústica de Campanhã*, **B** – *O mendigo Lapita*, **C** – *Paisagem – Abertura da Rua Alexandre Herculano*, **D** – *Jardim de Luxemburgo (estudo)*, **E** – *Aldeia de St. Sauves*, **F** – *Paisagem de St.Sauves*, **G** – *Cansada (cachopa de Capri)*, **H** – *Casas brancas de Capri*, **I** – *Cecília*, **J** – *Escadas de um pardieiro – Roma*, **K** – *Esperando o sucesso*, **L** – *Fachada de casa soterrada – Roma*, **M** – *Miragem de Nápoles*, **N** – *Mulher da água*, **O** – *Portão*, **P** – *Rapariga de Anacapri*, **Q** – *Rua de Roma*, **R** – *Senhora vestida de preto*, **S** – *Janela das persianas azuis*, **T** – *Paisagem de Anacapri*, **U** – *Rapariga deitada num tronco de árvore*, **V** – *Cais de Barcelona*, **W** – *Flores campestres*

[a] identified by Scanning electron microscopy/energy dispersive X-ray analysis.

In addition to pigment elucidation, the information collected by Raman microscopy can be used for other purposes including aiding artwork owners in assessing whether a painting or other potentially valuable item is authentic. Such an example would include Clark's investigation of Mauritian postage stamps.[20] He-Ne (632.8 nm) and Ar (514.5 nm) lasers were used to create Raman excitation in this investigation. Unlike the previous paper, however, the stamps experienced no preparation prior to being placed under the microscope objective. This is likely to be due to the stamps' value and would also suggest that the laser power was kept as low as possible.

In 1847, Mauritius became the first British Colony and the fifth country in the world to issue postage stamps. At this time there were only two denominations: one penny or two pence. These were known as the Barnard issue, after the person who first issued them. Approximately 500 of each were made; however, only 14 of the 1847 one penny stamps and 12 of the 1847 two pence stamps are known to still exist.[20] Several different stamps were also investigated using Raman microscopy including: 1848-59 post Barnard issues, 1858-62 Britannia undenominated issues and 1913 French reproductions of the 1847 Barnard stamps. Raman microscopy revealed that the genuine 1847 two pence stamp, its 1913 reproduction and the 1858-62 Britannia-type undenominated blue stamp all contained Prussian blue, which was assumed to be due to the limited selection of blue pigments at the time. It was found that the major difference between the original and the reproduction stamps was the presence of ultra-marine blue in the original. Two forgeries of the 1859 two pence Lapirot (post Barnard) stamps were also examined. These forgeries were produced using photogravure, a printmaking process, which allowed them to be easily distinguished from the originals because of a large amount of fluorescence caused by the glossy film which coats them. This resulted in obscured Raman signals from the pigments beneath. Other pigments identified include lead (II, IV) oxide, chrome yellow and mercury (II) sulfide. This example illustrates the potential of Raman microscopy to provide rapid and non-destructive means to assist in determining authenticity.

(a) 

(b) 


**Figure 1.3 Photograph of (a) used one penny and (b) unused two pence 1847 Mauritian stamps.[20]**

In addition to identifying pigments, Clark has employed Raman microscopy to identify variants of verdigris, a green-blue pigment.[21] Verdigris is used as a general term to refer to copper (II) acetate, $[Cu(CH_3COO)_2]_2{\cdot}Cu(OH)_2{\cdot}5H_2O$, and several similar compounds which only differ in their ratio of $[Cu(CH_3COO)_2]$ to $Cu(OH)_2$ and in their degrees of hydration. This variation results in differing coordination geometries of the copper(II) and hence the range of colours from green to blue. Therefore the goal of this research was to determine whether Raman microscopy could distinguish between the small differences in compound structure at either of two predetermined temperatures.

This project involved the measurements at two temperatures: room temperature and at 150 K. Room temperature samples were analysed using an argon (514.5 nm) laser to excite the sample and a power ranging from 0.04 - 0.4 mW. Measurements were able to be made at 150 K through use of a Linkam THMS 600 stage by placing the samples inside a quartz crucible inside a cold cell and cooled using liquid nitrogen. They were analysed using a 50x objective with a long working distance and 0.4 mW laser power. Each sample was synthesised via a different method and using different materials. Each spectrum collected was characterised by bands at *ca* 3010, 2985, 2935, 1550, 1445, 1415, 1350, 940, 220 and 140 $cm^{-1}$. Individual groups of compounds showed further characteristic bands. It is important to note that the Raman spectrum of each verdigris variant could be differentiated from one another based on the combination of peaks produced and their relative intensities. Results are presented in **table 1.2**. Therefore, not only can Raman microscopy be used to distinguish pigment from one another, but this paper demonstrates the ability (and potential use) of the technique for distinguishing between variants of the same colour.

**Table 1.2 Comparison of Raman shifts ($cm^{-1}$) of bands for nine variants of verdigris (labelled A, B, B2, B4, C, D2, D3-3, D* and D5-1) and a commercial sample of copper (II) acetate monohydrate at room temperature. Bands in italics have Raman shifts which are similar for all samples.[21]**

| A ($cm^{-1}$) | B ($cm^{-1}$) | B2 ($cm^{-1}$) | B4 ($cm^{-1}$) | C ($cm^{-1}$) | D2 ($cm^{-1}$) | D3-3 ($cm^{-1}$) | D* ($cm^{-1}$) | D5-1 ($cm^{-1}$) | Copper(II) acetate monohydrate ($cm^{-1}$) |
|---|---|---|---|---|---|---|---|---|---|
| 3611.5 | – | – | 3628 | – | 3610 | 3607 | 3612 | 3618.5 | – |
| – | – | – | – | – | – | – | – | 3609 | – |
| 3579 | – | – | – | – | 3578 | 3570 | 3578 | 3572 | – |
| – | – | – | – | – | – | 3552 | – | 3557 | – |
| 3517 | – | – | – | – | 3519 | 3515 | 3512 | 3510 | – |
| – | – | 3473 | 3430 | 3468 | – | 3453 | – | – | 3480 |
| – | – | 3374 | – | 3373 | – | 3380 | – | – | – |
| 3405 | – | – | – | – | 3403 | – | 3407.5 | 3411 | – |
| 3257 | – | – | – | 3267 | 3258 | 3277 | 3269 | – | – |
| *3020* | *3026* | *3031* | *3012* | *3022* | *3019* | *3020* | *3019* | *3004* | *3022* |
| – | – | *3022* | – | – | – | – | – | – | – |
| *2988* | *2990* | *2988.5* | *2984* | *2981* | *2987* | *2986* | *2995* | *2980* | *2982* |
| – | – | *2981* | – | – | – | – | *2986* | – | – |
| *2936* | *2938* | *2936.5* | *2934* | *2940* | *2935* | *2936* | 2934 | *2933* | *2939* |
| – | 2857 | 2860 | 2854 | 2857 | – | – | – | – | – |
| – | – | – | – | 2790.5 | – | – | – | – | – |
| 2164 | – | 2162 | – | 2698 | 2678 | – | – | – | – |
| *1552* | 1573 | *1531* | *1553* | *1533* | *1551* | *1574* | *1549* | *1566* | *1552* |
| – | *1550* | – | – | – | – | – | – | – | – |
| – | 1526 | – | – | – | – | – | – | – | – |
| *1439* | *1441* | *1441* | ~1450 | *1446* | *1436* | ~1445 | *1441* | 1477 | *1439* |
| – | – | – | – | *1438* | – | – | – | *1438* | – |
| *1429* | *1414* | *1417* | *1429* | *1414* | *1409* | *1421* | *1408* | *1422* | *1418* |
| *1409* | *1407* | – | – | *1409* | – | – | – | – | – |
| *1353* | *1351.5* | *1360* | *1348* | *1360* | *1354* | ~1343 | *1345* | *1347* | *1360* |
| *1342* | – | – | – | – | *1342* | – | – | – | – |
| – | – | 1052 | – | – | – | – | – | – | 1051 |
| *935* | *941* | *948* | *941* | *948* | *936.5* | *931* | *938* | *926* | *948* |
| – | *932* | *938* | *933* | *938* | – | – | – | – | *937* |
| – | – | 702 | – | 701 | 716 | – | – | – | 702 |
| 650 | 679 | 684 | 680 | 681 | 649.5 | 643 | 646 | – | 683 |
| 615 | 621 | 633 | 618 | – | 614 | – | – | – | – |
| – | 522 | – | 516 | – | – | – | – | – | – |
| 501 | 501 | – | – | – | 498.5 | – | 497 | – | – |
| 485.5 | – | – | – | – | 486 | 484 | – | – | – |
| 474 | 473.5 | – | 469 | 471 | ~450 | – | – | 469 | – |
| 403 | 398 | – | 383 | – | 402 | – | 401 | – | – |
| – | 386 | – | – | – | – | – | – | – | – |
| – | 330 | 319 | 315 | 320 | – | – | – | – | 321 |
| – | 315 | – | – | – | – | – | – | – | – |
| – | – | 298 | 276 | 298 | – | 283 | – | – | 295 |
| – | – | 251 | 267 | 251 | – | – | – | – | 251 |
| 214.5 | 232.5 | 231 | – | 231 | – | – | 229 | – | 231 |
| – | – | 225 | – | 223 | 221 | 219 | – | – | 224 |
| – | – | 212 | – | – | – | – | – | – | 212 |
| – | 198 | 179 | 199 | 185 | – | – | – | – | 178 |
| – | 183 | – | – | 178 | – | – | – | – | – |
| 159 | 164 | – | – | – | 157 | – | 156 | – | – |
| 144 | 141 | 140 | 132 | 141 | – | 139.5 | 141 | 139 | – |
| – | 135.5 | – | – | – | – | – | – | – | – |
| 112 | 121 | 124 | – | 125 | – | – | – | – | – |
| – | – | 117 | – | – | 111 | – | – | – | – |
| – | – | 107 | – | 103 | – | – | – | – | – |

It is evident that Clark's endeavours in the analysis of pigments using Raman microscopy have been extensive, and in many cases, successful. It suggests that the elucidation of pigments in 'Hope' is likely to be possible and demonstrates that minimal sample preparation is required (and in some cases preferred). Clark has provided a large foundation upon which Raman studies of pigments can be built. This thesis utilises the large amount of data provided by Clark to aid in identifying pigments from Edward Burne-Jones' 'Hope'.

## 1.4 Dairy samples

Milk plays an important role in the diet of mammals. For many mammals it is their only source of food during infancy. It contains all the nutrients required for the relatively rapid growth of their young. A summary of the content of milk from a typical cow can be found in **table 1.3**. This figure shows that other than water, bovine milk is mostly made up of fat (primarily as saturated, monounsaturated and polyunsaturated triacylglycerols), protein (80% casein, 20% whey protein) and lactose. [22, 23] The remainder of milk is known as ash, the oxides of minerals obtained after combustion.

**Table 1.3 Average component content of western bovine milk.[22]**

| Component | Average percentage of mass |
|---|---|
| Water | 86.6 |
| Fat | 4.1 |
| Protein | 3.6 |
| Lactose | 5 |
| Ash | 0.7 |

Cheese is an important food made from milk. The constituents vary due to the variable methods by which cheese can be made.[24] Currently, the western cheese making process begins with standardisation of the amount of fat and protein in liquid milk, which can be altered depending on the type of cheese which is desired. Minerals can also be added at this point depending on nutrient content desires (calcium, for example). The milk then goes through pasteurization which limits microbial numbers and enables the milk to have more uniform and controlled bacterial content. This milk is then converted to cheese curd through three procedures: acidification, coagulation and dehydration.[24] Acidification can be carried out

through either bacteriological production of lactic acid (fermentation of lactose by lactic acid bacteria) or the direct addition of lactic or hydrochloric acid. Coagulation refers to the coagulation of the casein within the milk and can be achieved by addition of rennet, a collection of enzymes found in mammalian stomachs capable of performing proteolysis. Coagulation can also be achieved through acidification to pH 4.6, or by acidification to higher pHs but with heating to 90 $^{o}$C. Dehydration is simply accomplished by cutting the curd, heating it and/or stirring it, thus expelling any unwanted whey. The 'last' step of the cheese making process is the salting process. However, during this process several different methods of treatment can be employed to change the final product, such as cheddaring or molding. Fresh cheese has then been made, yet further customisation of cheese can be attained through ripening. This is a highly variable process governed by the microflora present in cheese post-manufacture and can occur for any time between 3 weeks to 2 years, after which time a mature cheese is obtained.[24] **Figure 1.4** summarises this process.

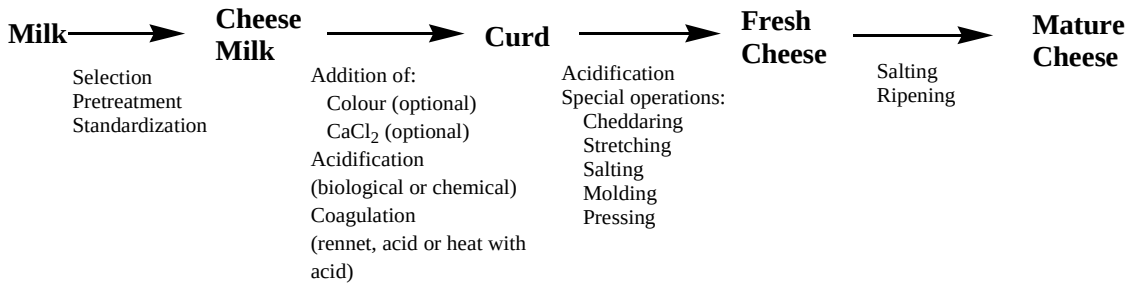



**Figure 1.4 A summary of the general process in which cheese is made.[24]**

Milk powder is an important food ingredient made through the dehydration of milk. Differences between milk powders are mostly limited to the different fat to protein ratio and mineral addition. Production of milk powder generally involves initial heat treatment of the milk which destroys pathogenic bacteria, deactivating certain enzymes such as lipase and to activate SH groups of β-lactoglobulin which increases resistance to oxidative change.[25] Evaporation of water then occurs. This is done by passing the milk through a system of about seven or eight effects, or devices which heat the milk under a partial vacuum. At this point, homogenisation can occur (but is not obligatory) in order to create a more consistent milk sample, usually through decreasing free fat content. Drying then occurs and can be achieved either by roller drying or spray drying. Roller drying may occur either under atmospheric

pressure or under vacuum. Spray drying is the more prominent method for drying. It is achieved through dispersing the milk into very fine droplets (atomisation) and exposing it to hot air flow. The dry powder then moves through a series of pipes and cyclone separator systems which discharges the air and collects the product.[25]

Compared with paint studies, studies of milk and its components using Raman spectroscopy are rare, and investigations of cheese even more so. In 2008 an incident in China where the adulteration of infant formula with melamine resulted in 6 deaths, and 294,000 affected infants with more than 50,000 being hospitalised.[26] Previously, techniques such as enzyme-linked immunosorbent assay (ELISA) and high performance liquid chromatography (HPLC) were used to detect melamine;[27] however, such techniques require time consuming preparation processes. Raman spectroscopy does not usually require such preparation and is quick and easy to use. This could make it an appealing alternative for studying dairy samples. A study has shown through use of the portable and compact Inspector Raman $^{TM}$ 1.3 (DeltaNu Inc., Laramie, WY, USA) melamine concentrations of 3, 5, and 10% (w/w) were able to be detected in milk powder using the characteristic peaks of the adulterant at 672 and 982 $cm^{-1}$. Through use of partial least squares (PLS) analysis, the detection limit of the technique was found to be 0.13% (w/w), which is lower than some amounts found in contaminated batches of infant formula (was found in amounts as high as 0.2563%).[28] Therefore such studies have begun to shed light upon the strength of Raman spectroscopy in the analysis of dairy products. Principal component analysis (PCA) and PLS analysis will be discussed often in this thesis. These techniques are discussed in **sections 4.3.1** and **4.5.1**. Further reading about these techniques can be found in reference 29.

As well as the possible presence of contaminants in dairy products, the components which make up the product itself have become the focus of many Raman studies. Fat and protein content are of interest due to the importance of such components in human nutrition. For example, fats are a concentrated source of energy with 37 kJ $g^{-1}$ compared with 16 kJ $g^{-1}$ from carbohydrates and 17 kJ $g^{-1}$ from protein.[30] Fat is an excellent source of energy and protein is both a source of energy and essential amino acids which are required for producing enzymes and for muscle growth.[30]

Quantification of fat in milk using Raman spectroscopy has already been the topic of prior research. Materny et al.[31] utilised Raman spectroscopy for this purpose. Materny used three different sample preparations: liquid milk in an open dish, dried milk on glass plates covered with aluminium foil, and liquid milk contained in quartz cuvettes. A spectral window of 800-3050 $cm^{-1}$ was observed. Through use of partial least squares analysis it was found that milk in the open dish and dried milk droplets allowed for an effective model for fat quantification. Liquid milk in cuvettes provided a poor quantification model. This was thought to be due to the quartz surface reducing the amount of incident and reflected light and due to the formation of a turbid, inhomogeneous coating on the inner surface of the cuvette. Quantification was achieved via correlation of concentration with relative band intensities and through construction of a plot of predicted fat content versus actual fat content.

The presentation of the milk sample is an important factor in its analysis. What is generally preferred is a fast, effective, consistent and cost effective method for preparation. However, fat from milk may be collected at various different concentrations ranging from diluted to pure anhydrous milk fat. McGoverin investigated the ability of Raman spectroscopy to predict solid fat content (SFC) and triacylglycerol (TG) class concentrations in anhydrous milk fat, which had previously been quantified by low resolution nuclear magnetic resonance (NMR) and high performance liquid chromatography (HPLC).[32] Raman spectra were collected at temperatures ranging from 0 - 50 $^{o}$C in 10 $^{o}$C increments (temperature of data collection: $TDC_0$, $TDC_{10}$…$TDC_{50}$). This was achieved using a water bath designed to permit *in situ* Raman measurements. As previously mentioned, there are two factors of interest: SFC and TG class content. There were 8 TG classes of interest including hard long saturates (HLS), hard short saturates (HSS), long monounsaturated (LM), short monounsaturated (SM), *c*-dienes (CD), *t*-monounsaturated (TM), *c*, *t*-dienes (CTD) and short saturated (SSat) which were all defined in McGoverin's paper.[32] Several different spectral transformation methods were utilised, including: baseline correction (BL), standard normal variate (SNV) and area normalisation (AN); baseline correction was sometimes coupled with either SNV or AN. SFC and TG class concentrations were treated separately using three different partial least squares methods: PLS1, PLS1 with modified Jack-knifing (PLS1-JK) and PLS2. Calibrations were subject to

full cross validation. McGoverin followed a full factorial design by Geir Flaten[33] for examining the six different TDCs, the seven different methods for preprocessing and three methods of analysis for all eight TG classes. The calibrations were characterised through root mean square error of cross validation (RMSECV) as discussed by Beebe[29] and the square root of RMSECV. RMSECV values were transformed by calculating the square root thus allowing an analysis of variance (ANOVA) to be applied to the resulting normal distribution which in turn allowed investigation of possible variation due to preprocessing technique, method analysis and TDC. The performance of each calibration treatment in predicting SFC levels or TG class concentrations was assessed through ranking of the mean square root of RMSECV and mean square error (MSE; explained by McGoverin to be the sum of the variance and bias).

Solid fat content of all levels were able to be reliably quantified from physically heterogeneous samples (at 20 $^{o}$C) using a single method of calibration involving TDC, preprocessing and PLS. However, McGoverin found that this was not possible for TG classes using a single method of calibration. Indeed, reliable prediction was found to be impossible for three TGs: SSat, SM and LM. Quantification of TG class concentrations was found to be most effective when samples were liquid. Raman spectroscopy has shown it can be used as a viable alternative to HPLC and NMR for quantification of SFC and TG classes in anhydrous milk fat.

Gallier studied milk fat globules (MFG) with respect to the ability of Raman microscopy to identify fingerprint characteristics of different sized MFGs and milk sources.[34] One particular advantage of Raman spectroscopy is that water does not strongly interfere with spectra.[35] This allows components in milk to be analysed in solution rather than requiring prior drying, thus removing the potential for variation from this source. Raw milk was collected from 2 breeds of cow, Jersey and Friesian-Jersey, centrifuged and the top cream layer was reconstituted to a 20% fat content using a phosphate buffered saline buffer with a pH of 7.2, then diluted ten-fold. The globules were considered ‘small’ (< 2 μm), ‘intermediate-sized’ (3-9 μm), or ‘large’ (> 10 μm). This study produced a comprehensive table of identified peaks found in the milk samples (**figure 1.5**) and several spectra of milk fat globules, one of which is also shown in

**figure 1.5**. The most notable peaks include the following: 1303 ($CH_2$ twisting), 1443 (saturated fatty acid $CH_2$ scissoring), 1654 (C=C) and 1744 $cm^{-1}$ (ester C=O). Each of the listed peaks are important in establishing the distribution of AMF within mixtures analysed in this thesis. The range of compounds and functional groups found produced distinguishable features between MFGs of different sizes and those from different breeds. One such feature is the presence of triglyceride, which is detectable in larger fat globules but not in the small fat globules. This can be seen by the presence of a peak at 1744 $cm^{-1}$ in larger MFGs. Variation in the amount of acyl chains (both liquid mobile and crystalline) was observable in milk fat globules of the same size, different size and those from different breeds of cow. Raman microscopy was also able to determine that carotenoids were present in higher quantity in FGs of Jersey cows than Friesian cows, as well as an increase in concentration with increasing size of FGs.

Main Raman peaks assignments[a].

| Origin | Wavenumber ($\nu$/ $cm^{-1}$) | Assignment/structural information obtained |
|---|---|---|
| Phospholipid headgroup | 845–895 | 846: Inositol residue<br>860: Phosphatidic acid<br>875: Choline |
| Carotenoids | 1006 | "Breathing" mode of ring compounds |
| Fatty acid | 1063–1083 | $\nu$(C–C) stretching<br>Symmetric phosphoryl stretching<br>*gauche* C–C |
| Saturated fatty acid | 1125–1133 | $\varphi$(C–C)<br>Out-of-phase C–C solid fat |
| Carotenoids | 1158 | $\nu$(C–C) |
| | 1213 | Antisymmetric phosphoryl stretching |
| PC, PI and PS | 1270 | C=C *cis* unsaturation<br>C–H in-plane bending of ethylene groups |
| PC, PI and PS | 1303 | $\delta$($CH_2$) twisting unsaturation |
| Cholesterol, | 1443 | Saturated fatty acid |
| PC, PI and PS | | $\delta$($CH_2$) scissoring |
| Carotenoids | 1526 | $\nu$(C=C) |
| PC, PI and PS | 1654 | $\nu$(C=C) *cis* unsaturation |
| Triacylglycerol | 1744 | Ester $\nu$(C=O) |
| Acyl chains | 2715–2726 | $\nu$(C–H) |
| Acyl chains in liquid state | 2850 | $CH_2$ symmetric stretching |
| Cholesterol, cholesterol ester | 2874 | $CH_3$ symmetric stretching |
| Acyl chains in crystalline state | 2885 ± 5 | Fermi resonance $CH_2$ stretching<br>$CH_2$ anti-symmetric stretching |
| Acyl chains | 2900–2910 | $\nu$(C-H) of acyl chain |
| Acyl chains | 2935 | $CH_3$ symmetric stretching |
| | 2960 | out-of-plane and $CH_3$ anti-symmetric stretching |
| | 3007 | *cis*-unsaturated =CH stretching |

[a] Adapted from Faiman & Larsson, 1976, Forrest, 1978, Bresson, Marssi & Khelifa, 2005, Krafft et al., 2005, Argov et al., 2008, and McGoverin et al., 2009.

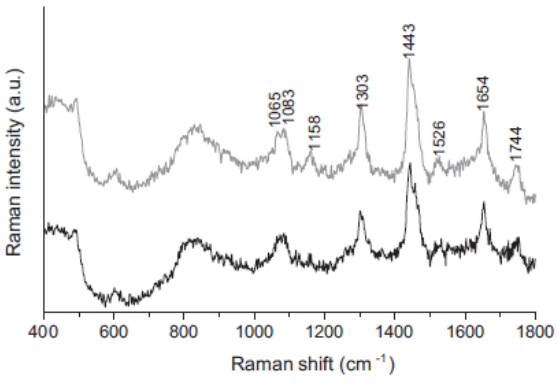



**Figure 1.5 A table showing assignment of MFG peaks (left) and a spectrum on the right comparing the 400-1800 $cm^{-1}$ region of MFGs collected from Jersey (grey) and Friesian-Jersey (black) cows.[34]**

McGoverin used Raman spectroscopy and chemometrics to identify and/or quantify the constituents in milk powder.[36] Principal components analysis (PCA) and visual analysis was used to identify outliers. Data was standard normal variate (SNV) pretreated before undergoing PLS analysis for each individual component. Validation variance scree plots, loading weights, regression coefficients, root mean squared error of prediction (RMSEP) and ratio of prediction to deviation (RPD) were all used to assess the reliability of results.

Using this array of information, McGoverin found Raman spectroscopy effective in quantifying fat and protein within milk powder, without the need for a specific method of sample preparation. This was achieved through use of regression coefficients which were associated with the relative characteristic peak heights of each component within the sample. During this experiment, the presence of a peak at 1087 $cm^{-1}$ in some samples suggested the presence of calcium carbonate, a known food additive.[37] This finding was confirmed through use of X-ray diffraction, suggesting a possible use for Raman spectroscopy to quantify calcium carbonate also.

There are relatively few Raman spectroscopy studies of milk / milk powder, yet the technique has proven very effective in the analysis of their components. Indeed, it has also shown potential in the quality control of milk powder.[38] This thesis is primarily focused on the application to cheese samples. Despite the large number of uses for Raman spectroscopy in the analysis of milk powder, the number of studies performed on cheese remains relatively low. One of the few investigations of cheese involved observing the effect of freezing and frozen storage on casein structure.[39]

The investigation involved three batches of cheese: A, B and C. Batch A was the control and experienced no freezing. Batch B was gradually frozen in a plate freezer, reducing the temperature to -35 $^{o}C$ at a rate of 7.5 $^{o}C\ h^{-1}$. Batch C was frozen quickly, reducing the temperature to -80 $^{o}C$ using a liquid nitrogen vapour freezer at a rate of 19.2 $^{o}C\ h^{-1}$. Two sets of results were obtained from both B and C. One set involved thawing for 48 h at 4 $^{o}C$ straight after freezing to investigate the effect of the freezing process. The other involved storing B

and C samples at -20 $^{o}$C for 4 months before thawing and then ripening. Afterward, the casein was extracted from each sample. Fourier-transform infrared (FTIR) and Raman spectroscopy were both used to analyse the samples. The results of this investigation suggest that both FTIR and Raman spectroscopy indicated that freezing increases the β-sheet and unordered structures within casein. However, being stored and ripened for 90 days showed a decrease in the difference in the amount of unordered structure between batch A and batches B and C. Furthermore, both techniques showed the proportion of unordered and turn structures was higher in cheese frozen in a plate freezer (batch B) than samples frozen more quickly using liquid nitrogen vapour.

Many of the papers discussed share similar goals to this thesis. McGoverin explored methods by which the components in milk and milk powder could be identified and quantified.[36] Materny used a different method and studied methods of quantifying fat in liquid milk samples.[31] Gallier, however, investigated the information which can be provided by the milk fat globules.[34] Furthermore, uses of Raman spectroscopy for monitoring milk powder adulteration are also being explored.[28] This thesis has the goal of distinguishing dairy samples and quantifying many of the constituents often found in dairy samples. However, these constituents are present in model systems. Milk powder samples contain larger calcite concentrations than those typically used for calcium fortification and the model cheese systems contain only two of four components often present in cheese. The four components of interest in this thesis are anhydrous milk fat (AMF), sucrose, cornstarch and rennet casein (type of casein present in cheese).

## 1.5 Thesis aims

The aim of this thesis is to explore the ability of Raman spectroscopic techniques to be used in the analysis of paint samples and certain dairy products. This involves:

- Investigating the ability of the Raman microscope to identify the pigments present in a painting by Edward Burne-Jones.
- Assessing the ability of FT-Raman spectroscopy together with various chemometric techniques to quantify the calcium carbonate content in spiked whole milk samples.

- Assessing the ability of Raman spectroscopy to distinguish model cheese mixtures based on component concentrations and the technique's ability to identify physical and chemical interactions between components.
- Investigating the ability of Raman microscopy to map the components of model cheese systems in two and three dimensions.

# Chapter 2 Experimental

## 2.1 Instrumentation

### 2.1.1 Raman microscope

Raman microscopy measurements used a Senterra confocal dispersive Raman microscope (Bruker Optics, Ettlingen, Germany). The variable Raman microscope parameters include (but are not limited to): laser wavelength (532 or 785 nm), laser power, objective (20x, 50x or 100x), aperture size, scan time and diffraction grating which affects resolution (there are two resolution settings: 3-5 $cm^{-1}$ or 9-18 $cm^{-1}$).

### 2.1.2 FT-Raman spectrometer

FT-Raman spectra were recorded using an Equinox 55 interferometer bench (Bruker Optics, Ettlingen, Germany) equipped with a FRA-106 Raman accessory and a D418-T liquid nitrogen cooled Ge detector. A 1064 nm Nd:YAG laser was used to generate Raman scattering. Two types of sample presentation were used; divet or cuvette. The cuvette was used for the linseed oil sample. All other samples were presented in divets. The horizontal objective was used to analyse the linseed oil sample. All other samples were analysed using the vertical objective. Both were in $180^{o}$ backscattering arrangements. Resolution was 4 $cm^{-1}$ for all samples. Laser power and number of scans were variable parameters.

### 2.1.3 Confocal laser scanning microscope (CLSM)

Samples were analysed in two different locations using two different confocal laser scanning microscopes. In Dunedin the microscope used was a Zeiss LSM710 (Carl Zeiss Inc., USA). In the Fonterra research facility in Palmerston North, the microscope used was a Zeiss LSM510 (Carl Zeiss Inc., USA). The variable parameters include (but are not limited to): excitation laser wavelength and the wavelength range of fluorescence bands collected.

## 2.2 Software

The Raman microscope was operated using OPUS 6.5 (Bruker Optics, Ettlingen, Germany). The FT-Raman spectrometer was operated using OPUS 5.5 (Bruker Optics, Ettlingen,

Germany). Remote manipulation of the spectra was accomplished using OPUS 6.5 and OriginPro 7.5 (OriginLab Corporation, Massachusetts, USA). The confocal laser scanning microscopes were operated using ZEN 2011 (Carl Zeiss Inc., USA). Remote manipulation of CLSM images was accomplished using Image J (National Institutes of Health, USA). Contour maps were manipulated using OriginPro 7.5, Adobe Photoshop CS 5.5 (Adobe Systems Inc., San Jose, U.S.A.) and Image-Pro Plus (version 7, Media Cybernetics, Maryland, USA). Standard normal variate (SNV) pre-processing, Principal component analysis (PCA) and partial least squares (PLS) analysis were performed using The Unscrambler X (CAMO Technologies Inc., Woodbridge, USA). Microsoft Office 2007 (Microscoft Corporation, Washington, USA) programs were used for a variety of purposes including data processing and presentation.

## 2.3 Paint, Skin and Oil Sample Analysis

Paint samples did not require preparation prior to being placed underneath the microscope objective. They were stored in transparent capsules (illustrated in **figure 3.1**) and, due to their small size, an attempt to move them onto a slide or other such platform was likely to have resulted in misplacing them. Therefore, they were analysed by Raman microscopy while inside the capsules. The skin sample was obtained from my own hand; near the fingernail. It was then placed on a quartz slide and analysed using the Raman microscope. The linseed oil sample required no prior preparation. Linseed oil was placed into a cuvette and analysed using the horizontal objective of the FT-Raman spectrometer.

### 2.3.1 Paint fleck 1 analysis using the Raman microscope

Fragment A was analysed using the 785 nm laser wavelength, 20x objective, 50 μm pinhole aperture, 50 mW power and 20 x 5 s integration time. Fragment B was analysed using the 785 nm laser wavelength, 20x objective, 50 μm pinhole aperture, 25 mW power and 110 x 5 s scan time. Fragment C was analysed using the 785 nm laser wavelength, 20x objective, 50 μm pinhole aperture, 25 mW power, 30 x 5 scan time and high resolution diffraction grating (3-5 $cm^{-1}$).

### 2.3.2 Paint fleck 2 analysis using the Raman microscope

This paint fleck had noticeably different pigments on opposite sides of the fleck. Blue and white pigments were found on one side and red and white pigments were found on the other. In order to analyse both sides, the capsule was lightly tapped which resulted in the fleck dropping to the other side of the capsule. This was repeated until the desired side was visible to the microscope. The Raman microscope parameters used to analyse all regions of paint fleck 2 were: 785 nm laser wavelength, 20x objective, 25 mW power, 50 μm pinhole aperture, 35 x 5 s scan time and high resolution diffraction grating (3-5 $cm^{-1}$).

### 2.3.3 Paint fleck 3 analysis using the Raman microscope

The Raman microscope parameters used to analyse all regions of paint fleck 3 were: 785 nm laser wavelength, 20x objective, 25 mW power, 50 μm pinhole aperture, 20 x 5 s scan time and high resolution diffraction grating (3-5 $cm^{-1}$).

### 2.3.4 Skin fleck analysis using the Raman microscope

The Raman microscope parameters used to analyse the fleck of skin were: 785 nm laser wavelength, 20x objective, 100 mW power, 50 x 1000 μm aperture, 10 x 5 s scan time and the low resolution diffraction grating (9-18 $cm^{-1}$).

### 2.3.5 Linseed oil analysis using the FT-Raman spectrometer

The FT-Raman spectrometer parameters used to analyse the linseed oil sample include: 1064 nm laser wavelength to generate Raman scattering, resolution was 4 $cm^{-1}$ for all samples, laser power was 300 mW and each sample was scanned 256 times. This sample was placed in a cuvette and analysed using the horizontal objective.

## 2.4 Analysis of Milk Powder Samples

Calcium carbonate was mixed with Pams® Instant Whole Milk Powder via geometric dilution using a mortar and pestle. The milk powder samples include 10%, 5%, 2.5%, 2%, 1.5%, 1%, and 0.5% (weight by weight; w/w) calcium carbonate. Three different sets of these concentrations were prepared and were referred to as sample sets (SS) 1-3. The process for creating these samples is illustrated in **table 1.1**. Each sample set was analysed on 3 separate

occasions, often referred to as either ‘repetitions’ or ‘days’ in Chapter 4. Samples were packed by hand into a small divet (approximately 10 mg) and analysed using the vertical objective of the FT-Raman spectrometer. Three spectra were taken for each sample on each occasion, rotating the divet between each spectrum. The FT-Raman spectrometer parameters used to analyse the spiked milk powder samples include: 1064 nm laser wavelength to generate Raman scattering, resolution was 4 $cm^{-1}$ for all samples, laser power was 120 mW and each sample was scanned 64 times. The resulting spectroscopic data was standard normal variate (SNV) pre-treated to remove signal offsets. The data was then processed using principal component analysis (PCA), partial least squares analysis (PLS) and univariate analysis. Standard normal variate pre-processing (SNV), PCA and PLS were performed using The Unscrambler X.

**Table 2.1 A table showing the process of creating each spiked milk powder sample. CC refers to calcium carbonate and MP refers to milk powder. Each mass was measured within ± 1 mg.**

| X | Mass of X added (g) | Mass of MP added (g) | Product name | Concentration of CC in product (w/w) |
|---|---|---|---|---|
| CC | 1.0 | 1.0 | Product 1 | 50% |
| Product 1 | 2.0 | 2.0 | Product 2 | 25% |
| Product 2 | 4.0 | 4.0 | Product 3 | 12.5% |
| Product 3 | 8.0 | 2.0 | Product 4 | 10% |
| Product 4 | 1.0 | 1.0 | Product 5 | 5% |
| Product 5 | 1.0 | 1.0 | Product 6 | 2.5% |
| Product 6 | 1.0 | 0.25 | Product 7 | 2% |
| Product 7 | 0.50 | 0.50 | Product 8 | 1% |
| Product 8 | 0.50 | 0.50 | Product 9 | 0.5% |
| Product 7 | 0.60 | 0.20 | Product 10 | 1.5% |

## 2.5 Cheese Sample Preparation

### 2.5.1 Preparation of binary mixture set A (BMA)

Binary mixtures were made up containing AMF (anhydrous milk fat) and one other component (component B). In each mixture component B was rennet casein, sucrose **or** cornstarch. Mixtures include 20%, 40%, 60% or 80% AMF (w/w) with the remainder of the mixture being component B, and were created via geometric dilutions. The term ‘component B’ is used to generally refer to rennet casein, sucrose or cornstarch. **Table 2.2** illustrates how BMA was prepared.

**Table 2.2 A table showing the process by which AMF/component B mixtures were made. The mixtures are referred to by the concentration of AMF they contain. 'B' refers to component B. Each mass was measured within ± 2 mg.**

| | | Step 1 | Step 2 | Step 3 | Step 4 |
|---|---|---|---|---|---|
| X | AMF concentration (w/w) | Mass of AMF added (mg) | Mass of B added (mg) | Mass of X added (mg) | Mass of X added (mg) |
| B | 20% | 100 | 100 | 200 | 100 |
| B | 40% | 200 | 200 | 100 | N/A |
| AMF | 60% | 200 | 200 | 100 | N/A |
| AMF | 80% | 100 | 100 | 200 | 100 |

### 2.5.2 Preparation of binary mixture set B (BMB)

These mixtures contain 1:1 proportions of AMF : component B. This was achieved by mixing 500 ± 2 mg of AMF with 500 ± 2 mg of component B for approximately 1 minute using a mortar and pestle.

### 2.5.3 Preparation of tertiary cheese component mixtures

388 ± 2 mg of previously prepared 1:1 mixtures (**section 2.5.2**) were mixed with 12 mg of trisodium citrate (TSC) using a mortar and pestle. The resulting mixture was then heated to 85 $^{o}$C for 8 minutes using a drying oven. The tertiary mixtures now containing 3% TSC had cooled to room temperature before being placed under the microscope.

## 2.6 Cheese Analysis

### 2.6.1 Distinguishing BMA mixtures using FT-Raman spectroscopy

All samples from binary mixture set A as well as pure AMF, RC, SUC and CS were analysed using FT-Raman spectroscopy in order to determine whether they could be distinguished based on the component concentrations they contained. The preparation of these mixtures can be found in **section 2.5.1**. The FT-Raman spectrometer parameters used to analyse binary mixture set A include: 1064 nm laser wavelength to generate Raman scattering, resolution was 4 cm$^{-1}$ for all samples, laser power was 400 mW and each sample was scanned 64 times. Three spectra were obtained for each mixture (the sample was rotated between each spectrum taken) and the results were pre-processed using SNV and processed using principal component analysis (PCA).

### 2.6.2 Identification of interactions within 1:1 binary mixtures

All samples of binary mixture set B (BMB) were analysed using Raman microscopy in order to discern whether chemical or physical interactions between components could be observed. The preparation of BMB samples can be found in **section 2.5.2**. Spectra were taken from each sample at 8-12 regions. A variety of locations within the sample were analysed and were intentionally selected to include a variety of rough and smooth regions. The Raman microscope parameters used to analyse all regions of binary mixture set B were: 785 nm laser wavelength, 50x objective, 100 mW power, 25 μm pinhole aperture, 6 x 10 s scan time and the high resolution diffraction grating (3-5 $cm^{-1}$).

### 2.6.3 Tertiary mixture analysis

The preparation of tertiary mixtures can be found in **section 2.5.3**. These mixtures were analysed using Raman microscopy in order to determine whether any chemical or physical interactions between components can be observed. Spectra were taken from each sample at 8-12 regions. A variety of locations within the sample were analysed and were intentionally selected to include a variety of rough and smooth regions. The Raman microscope parameters used to analyse all regions of the tertiary mixtures were: 785 nm laser wavelength, 50x objective, 100 mW power, 25 μm pinhole aperture, 6 x 10 s scan time and the low resolution diffraction grating (9-18 $cm^{-1}$). Spectra were compared to a spectrum of trisodium citrate (TSC) taken using the FT-Raman spectrometer in order to determine whether any regions in the tertiary mixtures included TSC. The FT-Raman spectrometer parameters used to analyse TSC include: 1064 nm laser wavelength to generate Raman scattering, resolution was 4 $cm^{-1}$ for all samples, laser power was 200 mW and each sample was scanned 256 times. Spectra from tertiary mixtures were also compared to various spectra of respective binary mixtures, depending on which components were present. For example, spectra of the AMF/CS/TSC mixture were compared to spectra of bi-component mixtures containing AMF and CS.

## 2.7 Producing Contour Maps

### 2.7.1 Producing 2D contour maps

Binary mixture set A (BMA) was used for creating contour maps. The preparation of these samples can be found in **section 2.5.1**. Samples were placed on a glass slide, fixed under the

objective of the Raman microscope and the stage was calibrated. Using OPUS 6.5, a 10 x 10 measurement point grid was placed over a region of interest. The dimensions of the grid were 130 µm x 100 µm along the X and Y axes, respectively. A spectrum was taken from each point on the grid. The Raman microscope parameters used to analyse all points of all contour maps were: 785 nm laser wavelength, 50x objective, 100 mW power, 50 µm pinhole aperture, 5 x 5 s scan time and the low resolution diffraction grating (9-18 $cm^{-1}$). The resulting spectra were then used to create a single MAP file. Each component has at least one Raman shift at which it produces a peak which is absent in the spectra of all other components. These peaks were chosen to indicate the presence of the components. For example, rennet casein produces a peak at 1003 $cm^{-1}$ whereas AMF does not give a peak in this region. Therefore in mixtures containing AMF and rennet casein, the 1003 $cm^{-1}$ peak was used to indicate the presence of rennet casein. Spectra of each of the cheese components can be found in **figure 2.1**. Using OPUS, these peaks were integrated above a baseline which was defined as a straight line at the base of the peak between the start and end points of the integration. Characteristic peaks were integrated at each point on the 10 x10 grid and the integration values were then copied from OPUS into OriginPro 7.5 (Origin) where a table was created, relating the coordinates of each point to the peak intensities. Using Origin, matrices were produced and from the matrices the contour maps were made. The resulting contour maps then had the threshold colours adjusted and the separate component maps were combined using Adobe Photoshop CS 5.5 (Adobe Systems Inc., San Jose, U.S.A.). The process by which the contour maps were made is illustrated in **figure 2.2**.

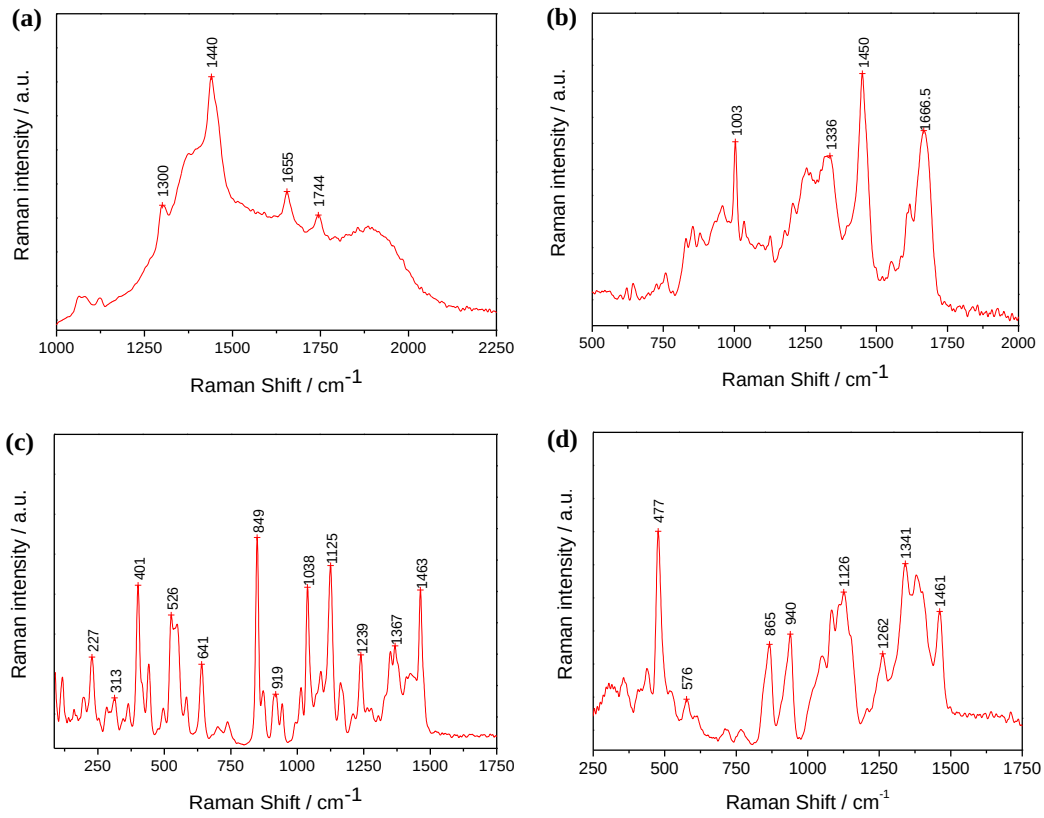



**Figure 2.1 Spectra of the pure cheese components used in the model cheese systems. (a) anhydrous milk fat, (b) rennet casein, (c) sucrose and (d) cornstarch.**

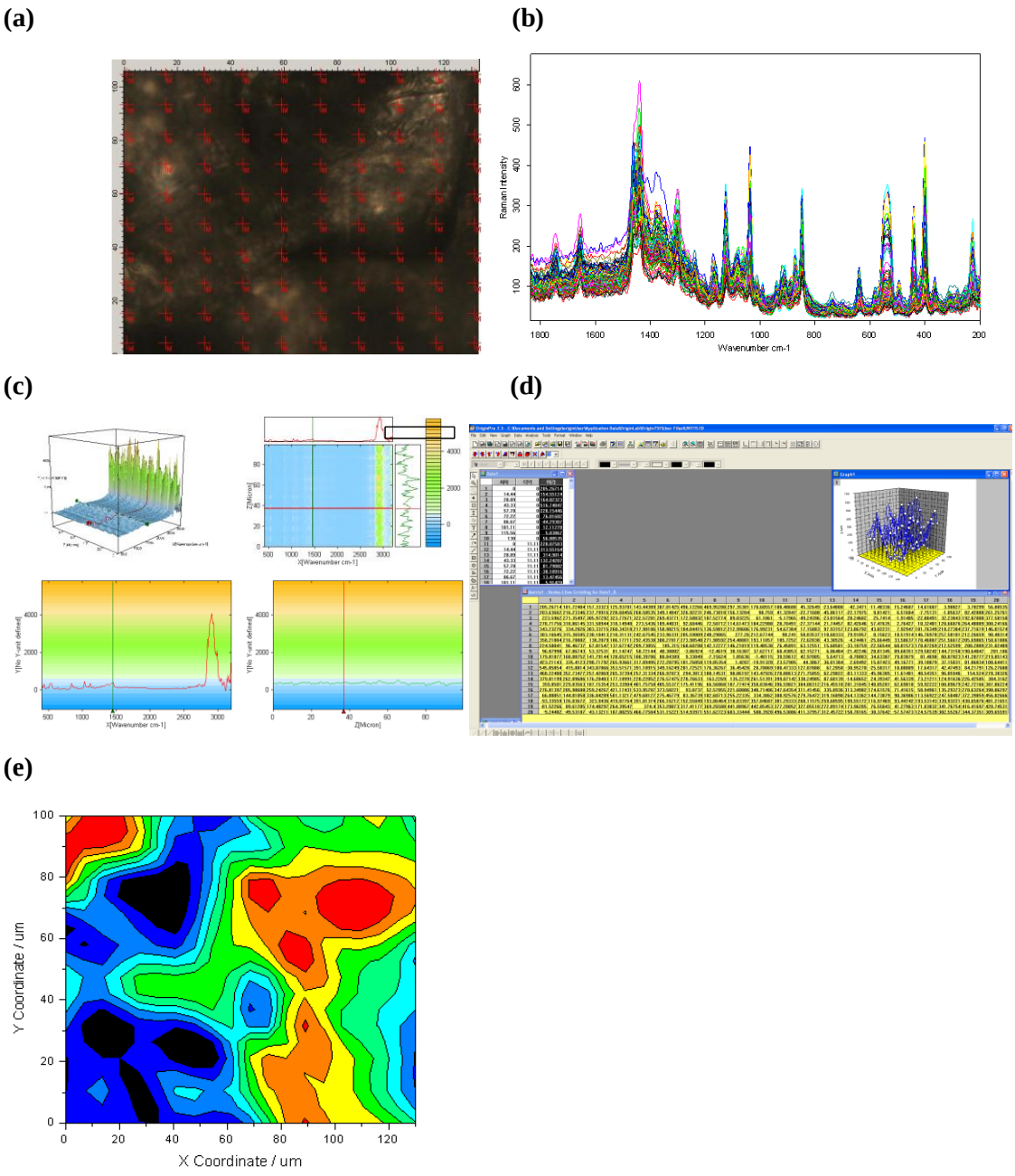



**Figure 2.2 Pictures illustrating the general process of creating the contour maps. (a) View from the microscope camera of the 100 point grid indicating where each spectrum was taken. (b) 100 spectra shown in OPUS 6.5 taken from the uniform 100 x 130 um grid. (c) The 100 spectra presented in the form of a MAP file. This image is present only to illustrate this step in creating the contour maps and units have not been defined at this point. (d) Picture of the OriginPro 7.5 software with: a table relating X, Y coordinates to peak integration values, the matrix created from the table of integration values, and an illustration of the relative intensities present in the matrix. (e) A contour map created from a matrix.**

### 2.7.2 Producing 3D contour maps

The samples containing 60% AMF and 40% of component B were used for creating the 3D contour maps (i.e. AMF60CS40, AMF60RC40 and AMF60SUC40). The method for preparation of these samples can be found in **section 2.5.1**. Creating the contour maps involves taking spectra at specific points in a 10 x 10 point grid of 130 um x 100 um dimensions; the same method performed in **section 2.7.1**. The Raman microscope parameters used to analyse all points of all contour maps were: 785 nm laser wavelength, 50x objective, 100 mW power, 50 µm pinhole aperture, 2 x 3 s scan time and the low resolution diffraction grating (9-18 $cm^{-1}$). Once one hundred spectra had been taken, the stage was lowered by 3.2 µm and one hundred more spectra were taken from a 10 x 10 point grid with the exact same X, Y coordinates as that of the grid placed 3.2 µm below it (see Appendix A). This process was repeated until 10 – 12 sets of 100 spectra were obtained. The contour maps were then manipulated in Origin to change the colours (e.g. to create a grey scale) and the threshold peak intensities used to indicate the relative quantity of a component in an area. For each set of contour maps the intensity thresholds were standardised to be the same for each contour map therein. To keep the method of picking a set of thresholds consistent, each contour map had the intensity thresholds synchronised with the fifth map from the bottom layer, which kept the same intensity thresholds calculated by Origin. Image-Pro Plus (version 7, Media Cybernetics, MD, USA) was used to combine the contour maps for the two components that are present in a mixture. Using the same software, the resulting map was thickened to give it 3dimensionality and each layer was stacked to give the final 3D contour map. An example can be seen in **figure 2.3**. Short video clips of the 3D contour maps being rotated are located on the supplementary DVD.

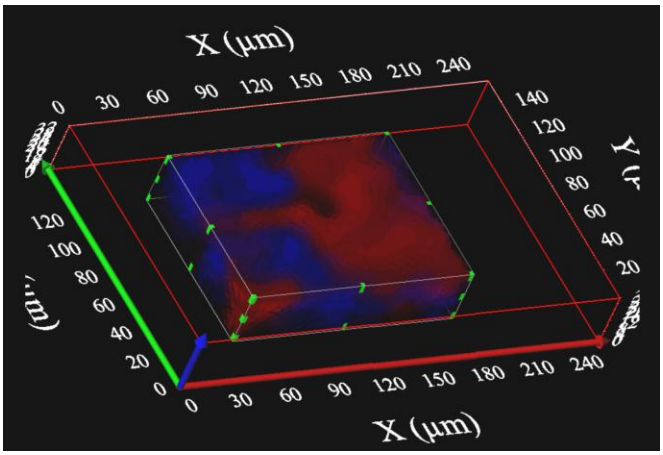



**Figure 2.3 An example of a 3D image (or 3D contour map) created using Image-Pro Plus. This image depicts the AMF60RC40 mixture.**

## 2.8 Producing Images using CLSM

Previously prepared AMF60RC40 (see **section 2.5.1**) sample was placed onto a glass slide and stained with 0.2% fast green and 0.5% nile red (weight per volume of citifluor; prepared by Elizabeth Nickless at the Fonterra Research Centre). The sample was left overnight in a refrigerator. In Dunedin, the microscope used was a Zeiss LSM710 (Carl Zeiss Inc., U.S.A.). The laser excitation wavelengths used were 488 and 633 nm and the bandpass filters used to collect sample emissions were 530-600 nm and 560-615 nm, respectively. At the Fonterra research facility in Palmerston North, the microscope used was a Zeiss LSM510 (Carl Zeiss Inc., U.S.A.). Laser wavelengths of 488 and 633 nm were used to excite the sample while fluorescence bands 575-615 nm and 650-730 nm were collected for the respective laser wavelengths.

# Chapter 3 Paint

## 3.1 Introduction

Artworks have been studied extensively using Raman microscopy.[16-21] The technique has proven to be effective at identifying pigments. This was demonstrated in a series of studies by Clark exemplified by reference 16. The purpose of Clark's study was to observe how Henrique Pousão's pallet changed over the time period the artworks were produced. Thirty seven compounds were identified. Most of these were pigments, but extenders and trace components were also included.

A series of studies were undertaken on the painting 'Hope' by Edward Burne-Jones to try to identify the materials present in paint flecks from this artwork. Three paint samples from 'Hope' were obtained from Jenny Sherman, the Dunedin Public Art Gallery conservator. The aim of this investigation was to identify the pigments and/or any binder that was used in the painting. Raman microscopy was chosen to analyse the samples because it is a non-destructive technique and can be applied to extremely small sample sizes. This is ideal because the paintings cannot be taken from the art gallery and excisions must be made to paintings which are preferred to be as small as possible. Spectra obtained from 'Hope' were compared to spectra of pigments obtained from the University College London (UCL) database.[40] It contains spectra from 56 pigments as well as tables containing their respective Raman shifts. The database was assembled by Clark with data collected from his own published investigations of pigments. This collection of data has been published collectively as a Raman spectroscopic library of pigments.[41] The UCL database website will be referred to in this thesis because it is the source of the literature spectra presented in this chapter.

## 3.2 Identification of pigments

Three samples were obtained from the Dunedin Public Art Gallery which were stored in three separate capsules. They were referred to as paint flecks 1, 2 and 3. Paint flecks 2 and 3 were found to be microscopically inhomogeneous therefore several areas of these flecks were

examined. These areas were identified by their optical appearance. **Figure 3.1** contains a photograph of the samples obtained from the Dunedin Public Art Gallery.



**Figure 3.1 A photograph of the capsule containing fleck 1 (left), capsule containing fleck 2 (middle), capsule containing fleck 3 (right). Paint flecks 2 and 3 are clearly visible; however, fleck 1 cannot be seen without using the microscope.**

### 3.2.1 Paint fleck 1

Examining the contents of the capsule containing paint fleck 1 illustrated some of the difficulties in analysis of very small samples. Several objects were located in this capsule using the microscope. The first of these was a white object (fragment A), approximately 20 x 25 μm and gave the spectrum shown in **figure 3.2a**. A darker object (fragment B) was found, it was approximately 40 x 50 μm and gave a different spectrum (**figure 3.2b**). The third object (fragment C) was white and considerably larger (approximately 100 x 150 μm). This fragment gave a spectrum shown in **figure 3.2c**; however, a picture was not obtained due to misplacement of the fragment.

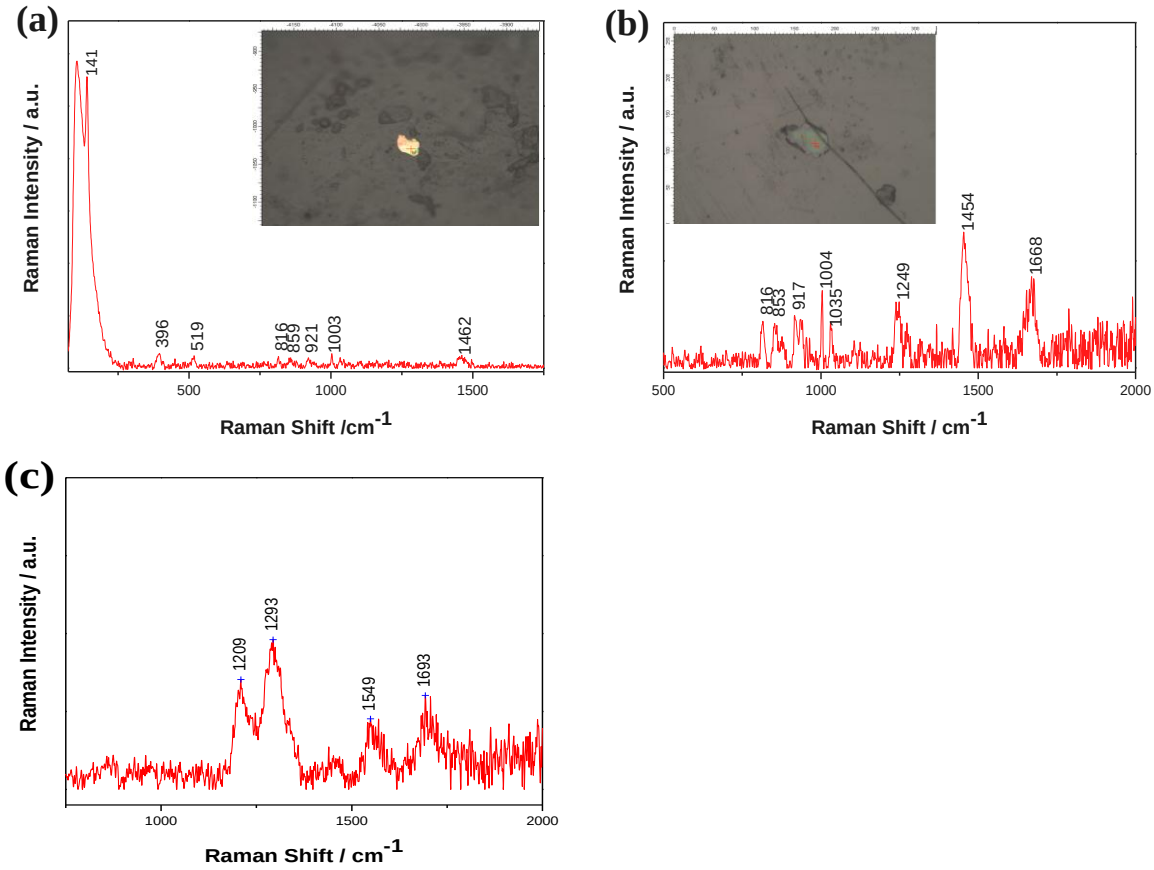



**Figure 3.2 Spectra of three different objects which were thought to be fragments of fleck 1: (a) fragment A, (b) fragment B and (c) fragment C. Pictures of each fragment accompany the spectra, except for fragment C which has no picture available. Parameters used for analysing fragment A include 785 nm laser wavelength, 20x objective, 50 µm aperture, 50 mW power and 20 x 5 s scan time. Parameters used for analysing fragment B include 785 nm laser wavelength, 20x objective, 50 µm aperture, 25 mW power and 110 x 5 s scan time. Parameters used for fragment C include 785 nm laser, 25 mW power, 20x objective, 50 µm aperture and 30 x 5 scan time. The width of each photograph is 320 µm.**

The spectrum of **Fragment A** is shown in **figure 3.2a**. This fragment produces a very strong band at 141 $cm^{-1}$ and several very weak bands at 396, 519, 816, 859, 921, 1003 and 1462 $cm^{-1}$. The bands present at 141, 396 and 519 $cm^{-1}$ are likely to be due to the presence of titanium white ($TiO_2$),[42] a pigment which is not found in the UCL database.[40] Peaks created by this pigment match in both relative intensity and Raman shift. The white colour is also indicative of the presence of anatase. None of the remaining peaks match any other pigment present in

the UCL database. However, the peak at 1003 $cm^{-1}$ is characteristic of phenylalanine, which would suggest that it may contain protein. The peak at 1462 $cm^{-1}$ is also similar in Raman shift (within ~10 $cm^{-1}$) to peaks recorded in literature for proteins and lipids.[43] This may suggest that this fragment is a fleck of skin with pigment attached. However, none of the remaining peaks could be attributed to lipid or protein; therefore this could not be confirmed.

The spectrum of **fragment B** is shown in **figure 3.2b** and is quite different from that of fragment A. It is characterised by bands at 816, 853, 917, 1004, 1035, 1249, 1454 and 1668 $cm^{-1}$. The peaks at 1004, 1249 and 1454 $cm^{-1}$ peaks match lipid and protein peaks found in skin.[43] The 1668 $cm^{-1}$ peak is also similar to the C=O stretching mode of an amide I, differing in Raman shift by ~7 $cm^{-1}$.[43] The spectrum taken from a fleck of skin, shown in **figure 3.3**, indicates that the spectrum of fragment B shares peaks at 853, 1004, 1035 and 1454 $cm^{-1}$ with that of skin. They also produce similar peaks around 1660 $cm^{-1}$; although they differ by 13 $cm^{-1}$. These results indicate that it is likely that fragment B is a fleck of skin.

The spectra of fragment B also share some similar characteristics with oil-like substances[44] which are sometimes used as media, the substances used to carry the pigments.[14] **Table 3.1** contains a table illustrating that fragment B peaks at 1004, 1035, 1454 and 1668 $cm^{-1}$ are similar to oil peaks. Purified linseed oil, which is often used as a medium, was provided by the Dunedin Public Art Gallery for comparison with the spectrum of fragment B. The spectrum of linseed oil is shown in **figure 3.4**. Both linseed oil and fragment B produced peaks around 1004 $cm^{-1}$. However, both the 1454 and 1668 $cm^{-1}$ bands differ considerably from the respective linseed oil bands in these regions. This suggests that it is unlikely that peaks in fragment B are caused by the presence of this oil medium. However, there are many more oils and oil mixtures which could have been used as a medium in this painting.[14] Nevertheless, these results suggest that fragment B is likely to be a fleck of skin.

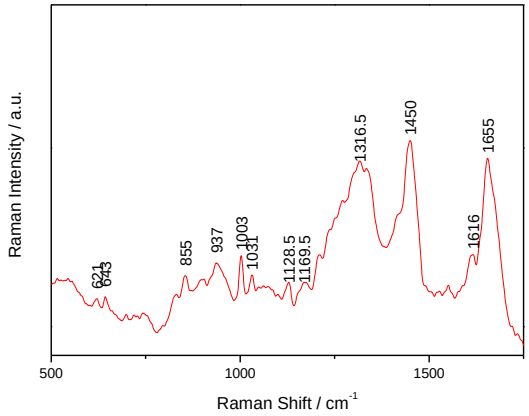



**Figure 3.3 A spectrum of a fleck of skin from my own hand. Parameters used for analysing the fleck of skin include 785 nm laser wavelength, 20x objective, 50 x 1000 µm aperture, 100 mW power and 10 x 5 s scan time and the low resolution diffraction grating.**

| **Raman Shift ($cm^{-1}$)** | **Molecule** | **Group** | **Vibration*** |
|---|---|---|---|
| 3015 | RCH=CHR | =C-H | ν asymmetric |
| 2970 | $-CH_3$ | C-H | ν asymmetric |
| 2940 | $-CH_2$ | C-H | ν asymmetric |
| 2900 | $-CH_3$ | C-H | ν symmetric |
| 2860 | $-CH_2$ | C-H | ν symmetric |
| 1750 | RC=OOR | C=O | ν |
| 1670 | *trans* RCH=CHR | C=C | ν |
| 1660 | *cis* RCH=CHR | C=C | ν |
| 1445 | $-CH_2$ | C-H | δ |
| 1310 | $-CH_2$ | C-H | δ |
| 1275 | *cis* RCH=CHR | =C-H | δ |
| 1000-1100 | $-(CH_2)_n-$ | C-C | ν |
| 800-900 | $-(CH_2)_n-$ | C-C | ν |

* ν, stretching; δ, deformation.

**Table 3.1 A summary of Raman shifts obtained from edible oils.[44] These are general values found in edible oils and were calculated using an average of values obtained from several different oils, one of which is Linseed oil.**

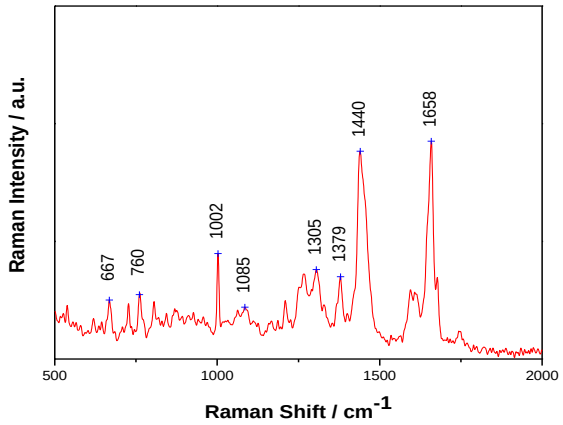



**Figure 3.4 A spectrum of Linseed oil taken using the FT-Raman spectrometer.**

The spectrum of **Fragment C** is shown in **figure 3.2c**. This fragment was white and similar in appearance to paint flecks 2 and 3. It was approximately 100 x 150 μm in size. Several spectra were taken from this fragment yet the spectrum shown contains the most well defined bands obtained. Bands are present at 1209, 1293, 1549 and 1693 $cm^{-1}$. These peaks were very weak and do not match any of the pigments on the UCL database.[40] The fragment was adhered to the inside of the capsule and in the process of turning the fragment over, its location was lost. Because of this, the pigments of paint fleck 1 could not be conclusively determined.

Considering the relative sizes of the fragments, it is possible that fragments A and B were simply contaminants. Spectral evidence suggests that fragment B is likely to be a fleck of skin, while fragment A may have been an actual fragment from fleck 1, or a fleck of skin with titanium white on it. The latter is not unlikely since the samples were prepared using bare hands in the Dunedin Public Art Gallery workshop area. Due to the lack of conclusive evidence, the elucidation of pigments from paint fleck 1 could not be achieved.

### 3.2.2 Paint fleck 2

Chronologically, this was the first fleck of paint to be analysed due to the elusive nature of fleck 1. Initial inspection of this fleck revealed that several different pigments were likely to

be present. Therefore, regions of this fleck were divided into three distinct categories based on their optical appearance; blue regions, white regions, and red regions.

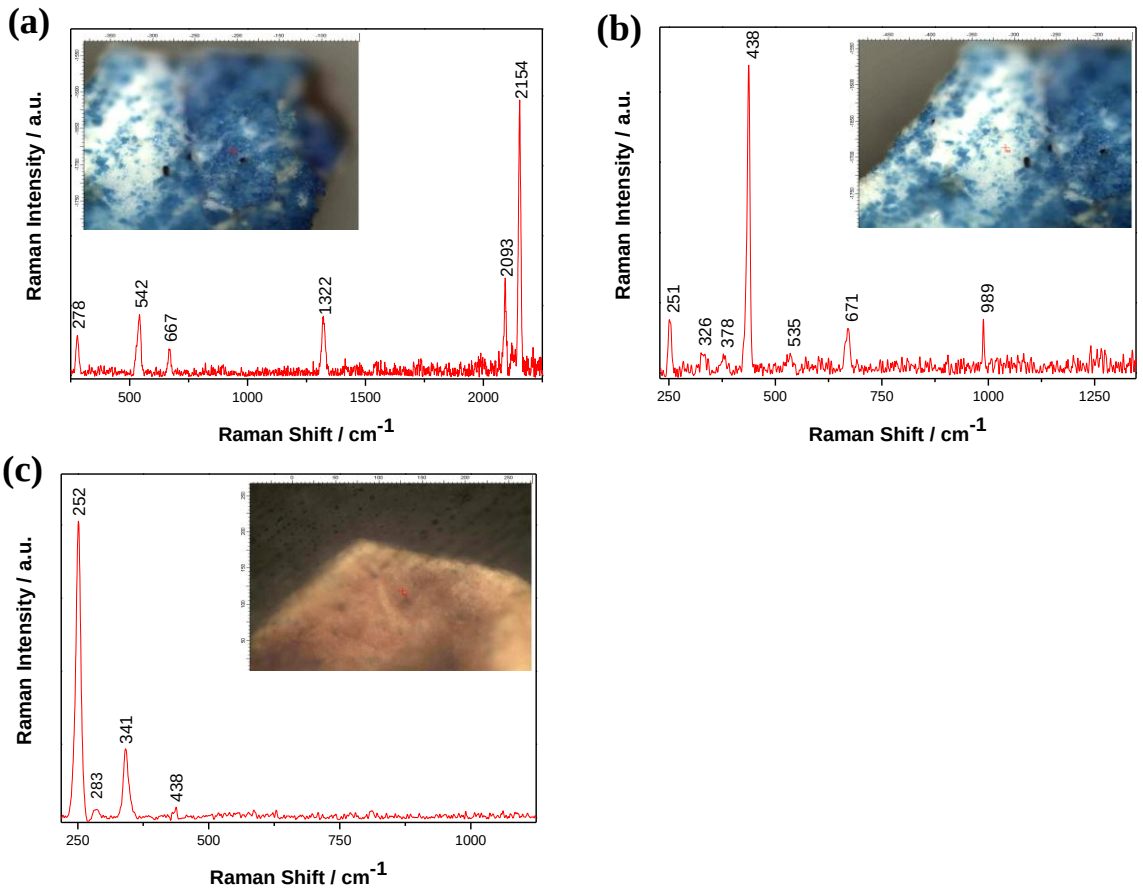



**Figure 3.5 Three spectra (and corresponding pictures) of three different regions of paint fleck 2: (a) a blue region, (b) a white region and (c) the red region on the opposite side of the fleck. The parameters used to analyse this fleck include 20x objective, 785 nm laser, 25 mW power, a 50 µm pinhole aperture, with a 35 x 5 s scan time. The width of each photograph is 320 µm.**

The spectrum from the **blue region** of fleck 2 can be found in **figure 3.5a** and is comparable to those found for cerulean and Prussian blue already reported in the literature.[40] A medium sized peak at 542 $cm^{-1}$ and a weak peak at 667 $cm^{-1}$ resemble two of the three peaks which could be seen on a spectrum of cerulean blue. The weak peak at 278 $cm^{-1}$, the medium peak at 542 $cm^{-1}$, the medium peak at 2093 $cm^{-1}$ and the very strong peak at 2154 $cm^{-1}$ are indicative of the presence of Prussian blue. **Figure 3.6** contains literature spectra of both cerulean and Prussian blue. These results suggest the paint is made from a mixture of cobalt(II) stannate ($CoO.nSnO_2$), and iron(III) hexa-cyanoferrate(II) ($Fe_4[Fe(CN)_6]_3.14\text{-}16H_2O$).

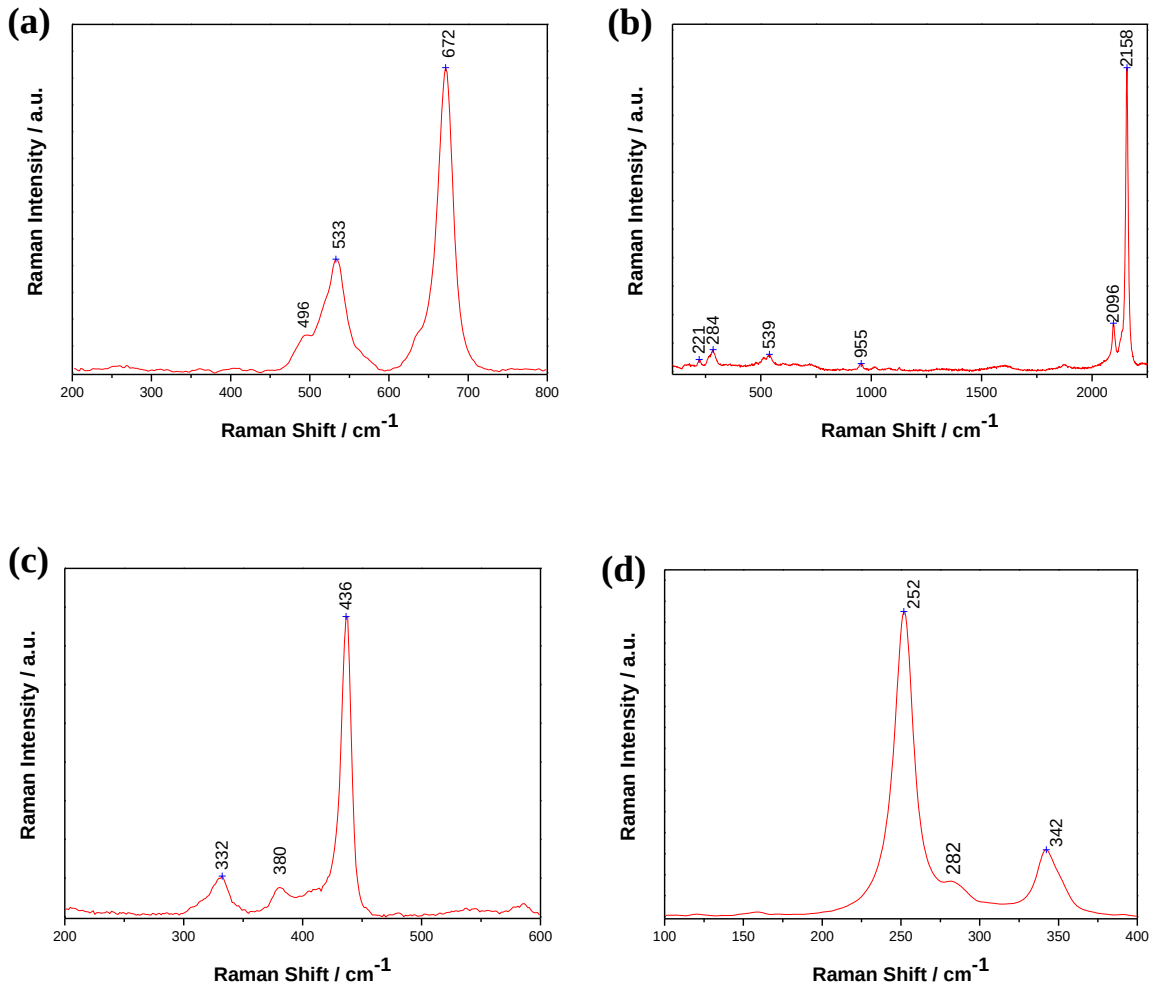



**Figure 3.6 Four literature spectra obtained from the UCL database: (a) cerulean blue, (b) Prussian blue, (c) zinc white and (d) vermilion.[40] Cerulean blue, Prussian blue and zinc white spectra were obtained using a 514.5 nm excitation wavelength. The spectrum of vermilion was obtained using an excitation wavelength of 632.8 nm.**

The spectrum of the **white region** (**figure 3.5b**) gives a weak peak at 326 $cm^{-1}$, a weak peak at 378 $cm^{-1}$ and a strong peak at 438 $cm^{-1}$. All three of these Raman shifts match literature values for the white pigment, zinc white (see **figure 3.6c**). Not only this, the relative peak intensities match also. This gives a strong indication that the paint used is composed of zinc oxide (ZnO). The source of the other peaks present is discussed later in **section 3.3**.

The spectrum of the **red region** (**figure 3.5c**) gives three significant peaks, a very strong peak at 252 $cm^{-1}$, a weak peak at 283 $cm^{-1}$ and medium sized peak at 341 $cm^{-1}$. These three peaks all closely resemble the three peaks for vermilion (see **figure 3.6d**) in both wavenumber and relative intensity. This gives a strong indication that this paint is made up of mercury (II) sulfide, (HgS).

### 3.2.3 Paint fleck 3

This sample was also found to be microscopically inhomogeneous which suggested that several pigments were likely to be present. Therefore, regions of this fleck were separated into three categories: light blue, dark blue and green.

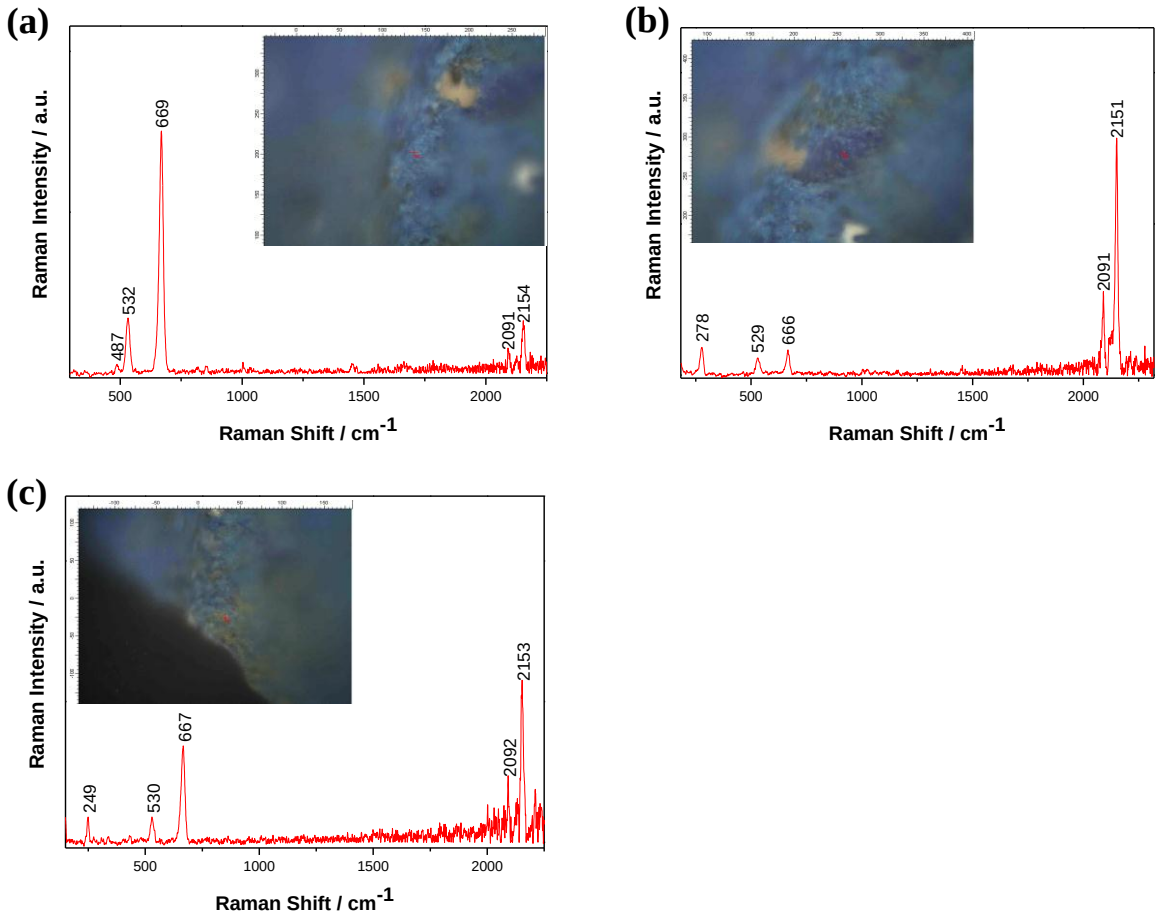



**Figure 3.7 Three spectra (and corresponding pictures) of three different regions of paint fleck 3 including: (a) a light blue region, (b) a dark blue region and (c) a green region. The parameters used to analyse fleck 3 were: 20x objective, 785 nm wavelength laser, 50 mW power, a 50 µm pinhole aperture, with a 20 x 5 s scan time. The width of each photograph is 320 µm.**

The **light blue region** (**figure 3.7a**) of paint fleck 3 gives a partially obscured weak peak at 487 $cm^{-1}$, a strong peak at 532 $cm^{-1}$, and a very strong peak 669 $cm^{-1}$. The Raman shifts and the relative intensities each match three peaks which are indicative for cerulean blue, indicating that this region is primarily composed of cobalt (II) stannate ($CoO.nSnO_2$). This region also gave peaks at 2091 and 2154 $cm^{-1}$ which are indicative of Prussian blue (refer to the literature spectra in **figure 3.6**).

The **dark blue region** (**figure 3.7b**) is characterised by: a medium peak at 278 $cm^{-1}$, a medium peak at 529 $cm^{-1}$, a weak peak at 2091 $cm^{-1}$ and a medium peak at 2151 $cm^{-1}$. The Raman shifts of these peaks resemble those seen in literature for Prussian blue. The presence of Prussian blue would suggest that it is made from iron(III) hexa-cyanoferrate(II) ($Fe_4[Fe(CN)_6]_3.14\text{-}16H_2O$. As seen in **figure 3.6a**, the peak at 669 $cm^{-1}$ can be explained by the presence of cerulean blue amongst the Prussian blue.

The spectrum of the **green region** (**figure 3.7c**) is of particular interest because it indicates that the pigment is simply a combination of Prussian blue and cerulean blue. The only peak which cannot be explained easily is the peak located at 249 $cm^{-1}$ which does not suggest the presence of any green pigment on its own. This peak is possibly caused by the presence of trace amounts of vermilion, a pigment which was abundant on the opposite side of paint fleck 2 which produces a very strong peak at this Raman shift. The obvious discrepancy between the apparent colour and the spectrum was vexing; however, after several spectra were taken to explore this phenomenon (not shown), all data suggested the presence of cerulean and Prussian blue.

## 3.3 Identification of binder

The rationale used for identifying binders was to locate all of the peaks which could not be assigned to pigments and determine whether there was enough information to classify which binder was used. After careful comparison of the paint flecks, no significant pattern in the content of the paint was found. The most common peaks to appear are those located between 660 and 670$cm^{-1}$ and those located around 2150 $cm^{-1}$ which can both be explained by different blue pigments. Cerulean blue gives a peak at 671 $cm^{-1}$ and Prussian blue gives a peak at 2158

$cm^{-1}$. The spectrum of the white region of paint fleck 2 (**figure 3.5b**) and the spectrum of the green region of paint fleck 3 (**figure 3.7c**) contain a similar peak around 250 $cm^{-1}$. This is possibly due to the presence of trace amounts of vermilion which has a strong band at 252 $cm^{-1}$ (**figure 3.6d**). The peak present at 438 $cm^{-1}$ in the red region is likely to be due the presence of some zinc white. The peak at 1322 $cm^{-1}$ on fleck 2 could not be assigned to any particular pigment and may indicate the presence of another compound but this peak on its own does not enable identification. Neither the pigment nor binder was able to be identified from paint fleck 1 because this sample was not conclusively located within the capsule. Therefore binder could not be identified in any of the spectra obtained.

## 3.4 Conclusions

Raman microscopy has proven effective in identifying pigments in 2 of the 3 samples provided by the Dunedin Public Art Gallery. Paint fleck 2 was found to contain cerulean blue, Prussian blue, zinc white and vermilion. Paint fleck 3 was found to contain cerulean blue and Prussian blue. Paint fleck 1 was not able to have its pigments identified. Not necessarily because of limitations to Raman spectroscopy but because of the difficulty of manoeuvring extremely small samples. Titanium white was able to be identified within the sample capsule, however, whether or not this came from paint fleck 1 or arrived through contamination is unknown. The binder used also remains unknown.

# Chapter 4 Milk powder

## 4.1 Introduction

Dried milk powder has several advantages over its liquid counterpart, namely the convenience of having a longer shelf-life and the ability to give a larger quantity of milk per unit of mass (assuming the availability of a source of potable water). Components of milk powder include: 25.2% protein, 38.1% carbohydrates, 26.2% fat, 7% ash and 3.5% water.[25] These proportions are likely to differ depending on the manufacturer of the product. Often, the nutrient proportions of such products are of interest to consumers which in turn, leads to the production of foodstuffs with altered nutrient content. Two examples would be reduced fat milk and milk containing a larger amount of calcium. Calcite is a common food additive used to increase calcium content of various foods, including milk powder, in order to maintain a healthy skeletal system, cellular metabolism and enzyme activation.[45, 46] The aim of this project is to see if Raman spectroscopy could be used to identify milk powder samples which have been spiked with calcium carbonate (calcite) and, if so, whether the calcite content can be quantified. This thesis employs several techniques to achieve this. Standard normal variate (SNV) pre-processing is utilised to reduce the effect of baseline differences between spectra. Principal component analysis (PCA) is used to differentiate between milk powder samples with differing calcite concentrations. Finally, univariate analysis and partial least squares (PLS) analysis are compared with respect to their efficacy for quantifying calcite concentrations in spiked milk powder. The samples analysed are separated into nine categories: each sample belongs to one of three sample sets (SS1, SS2 or SS3) and is also labelled repetition 1, 2 or 3 ('Rep 1, 2 or 3'). The 'sample set' label denotes that each of the samples was prepared on three separate occasions and the 'Rep' number indicates that each sample set was analysed on three different occasions. Literature has indicated that milk powder protein and fat can indeed be identified and effectively quantified using Raman spectroscopy.[36] This was achieved through use of a combination of Raman spectroscopy, X-ray diffraction, PLS analysis and PCA. The ability of Raman spectroscopy to identify the presence of crystalline constituents, lactose and calcite was also noted.[36]

## 4.2 Standard Normal Variate (SNV) Pre-Processing

Standard normal variate (SNV) pre-processing is a method by which the effects of baseline offset between spectra can be minimised. This is achieved by defining the average intensity as zero and scaling the total variance such that it is equal to one for each spectrum.[9] Each individual spectrum is processed using **equation 4.1** where the corrected intensity for spectrum ***i*** at wavenumber ***j*** ($x_{ij,SNV}$) is equal to the difference between the uncorrected intensity ($x_{ij}$) and the average intensity of the spectrum ($\overline{x}_i$) divided by the standard deviation of the intensity values from the same spectrum ($\sigma(x_i)$). This method is useful because it removes the baseline spectroscopic differences possibly caused by slight differences in sample presentation (such as sample surface layer morphology) and laser focus. Therefore, use of PCA or PLS after SNV pre-processing allows spectroscopic variation to be analysed without the data being confounded by baseline offset.

$$x_{ij,SNV} = \frac{x_{ij} - \overline{x_i}}{\sigma(x_i)}$$

**Eq. 4.1**

## 4.3 Principal Component Analysis (PCA)

### 4.3.1 Introduction to the technique

PCA was accomplished via The Unscrambler X software, using the NIPALS algorithm. The purpose of this technique is to identify the subtle relationships within data which contains a large number of variables. To do this, the original variables within the dataset are rotated and transformed into new variables which lie in planes of maximum variance. This is illustrated in **figure 4.1**. The new variables are known as principal components (PCs), latent variables or factors.[9, 29] Principal components are orthogonal to each other which ensures that each of the new variables are unrelated.[29] PCs are calculated in order of maximum variance. This means that PC1 represents the largest amount of variance, PC2 represents the second most, PC3 represents the third most, etc.

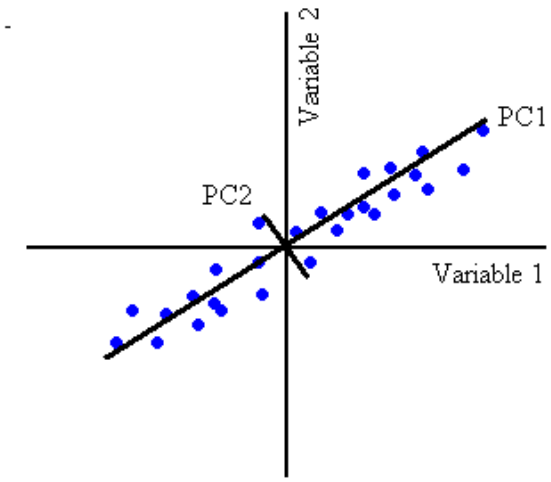



**Figure 4.1 A row plot of data in a two-variable system illustrating the relationship between PCs and the original variables.**[29]

The optimal number of PCs to use can be estimated through analysis of the amount of variance which each PC explains. In ideal situations the majority, if not all, variance should be explained by as few PCs as possible, which is usually 2 or 3 PCs. The major inflection point (the point at which the greatest change in slope occurs) within the explained variance or residual variance plot (Scree plots) can be considered to indicate the appropriate number to use.[9] The relationship between the old and new variables is demonstrated by the loadings plot. This plot essentially illustrates the contribution of the old variables, such as the Raman intensities at differing Raman shifts in spectra, to the principal component. This allows the regions of the spectra which account for the greatest amount of variance (largest contribution) to be distinguished from noise, the spectroscopic structures which are not relevant to the structure of the PC. Each analysed sample is assigned a score with respect to a particular PC. By plotting two PCs against each other, samples will be separated based on their scores for each PC. This plot is referred to as the scores plot and is the method by which the subtle differences in spectra due to differing constituent concentrations can be used to distinguish the samples.

### 4.3.2 Results and Discussion

The spectrum of pure calcium carbonate produces a very strong peak at 1085 $cm^{-1}$ (**figure 4.2**). This peak is easy to identify and is used for determining calcite presence in literature

also.[36] The 1085 $cm^{-1}$ is also quite apparent in the spectrum of the spiked milk powder. Hence it is the intensity of the peak located in this region which is of interest. The spectra of an array of milk powder samples spiked with differing amounts of calcium carbonate are illustrated in **figure 4.3**.

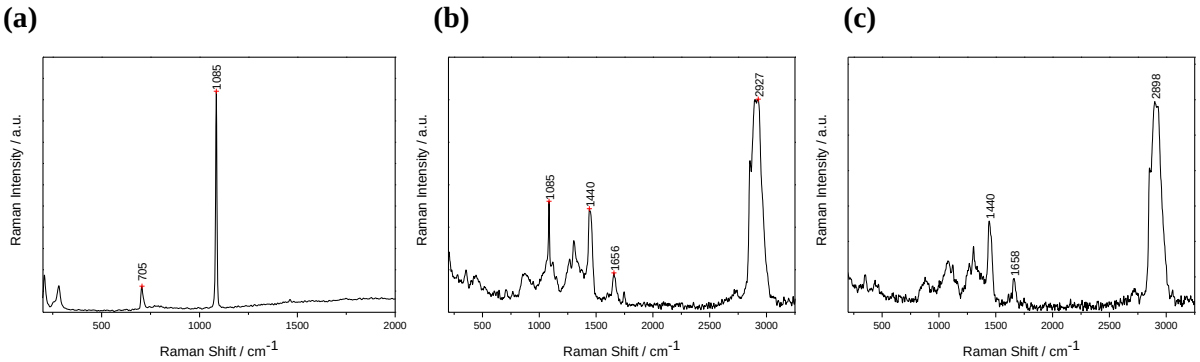



**Figure 4.2 Three spectra of different samples including: (a) pure calcium carbonate, (b) milk powder spiked with calcium carbonate and (c) pure milk powder.**

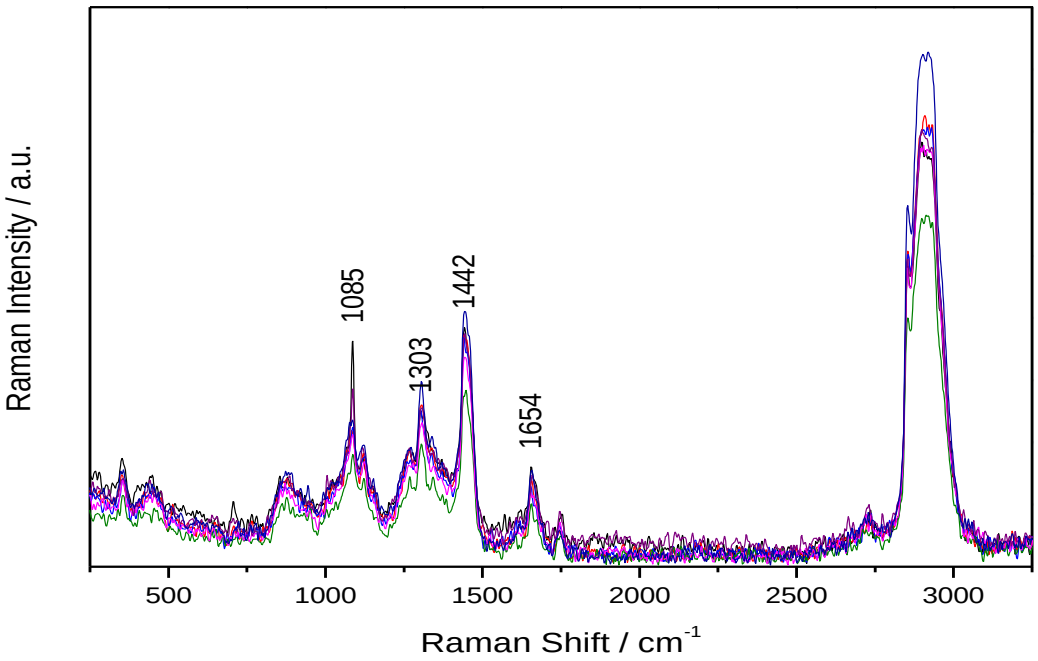



**Figure 4.3 A range of milk powder spectra, containing differing amounts of calcium carbonate.**

Due to the most prominent peak of calcium carbonate being located at 1085 $cm^{-1}$, only the region from 1178 $cm^{-1}$ to 987 $cm^{-1}$ was used for standard normal variate modification and principal component analysis. This ensures that the regions of the spectrum which are not

affected by the presence of calcium carbonate do not influence the principal component analysis with the natural variation which occurs between spectra. The analysed region is illustrated by the loadings plot in **figure 4.4**. The majority of variance can be seen to occur at approximately 1085 $cm^{-1}$.

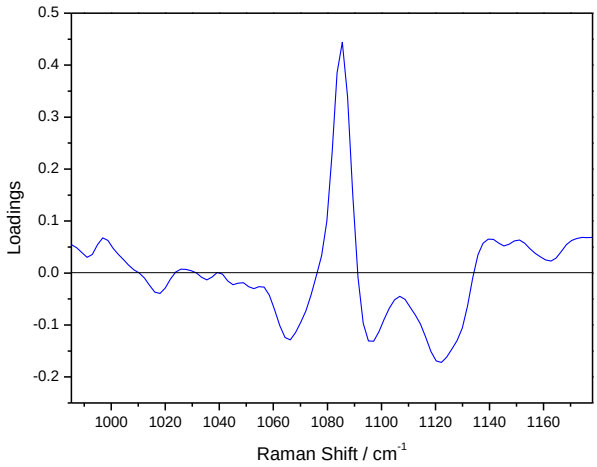



**Figure 4.4 Loadings plot describing the variance between 1178 and 987 $cm^{-1}$. Most of the variance occurs at 1085 $cm^{-1}$.**

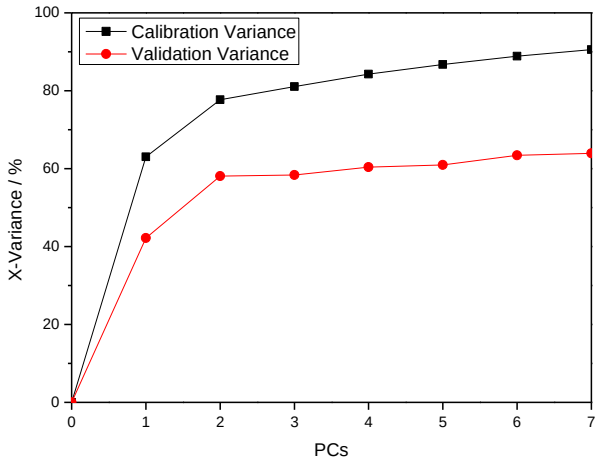



**Figure 4.5 A plot of the amount of variance explained by the principal components (SS1 repetition 1). The black line represents the calibrated variance and the red line represents the validated variance. The plot suggests that 78% of the variance in the spectra is explained by PC1 and PC2.**

**Figure 4.5** contains an explained variance plot constructed using 7 principal components. The black plotted line corresponds to the calibration variance. The red line refers to the validation

variance, which in this case was calculated using the full cross validation method. This plot can be utilised to establish the number of PCs which are useful for explaining the variance within a dataset. The optimum number of PCs to use can be considered to be the point of inflection of the plot,[9] which in this case is at 2 PCs. At this point, 78% of the variance is explained and the remainder of variance is evenly distributed between a large number of further PCs, each of which explain little of the remaining variance. This suggests that two PCs should be considered the optimal number to use to explain the variance within the data. The majority of sample set repetitions showed a similar trend, indicating that two PCs should be used. However, sample set 3 (SS3) was different and had the major inflection point located at PC1 with PC2 explaining little more variance than PC3. This suggests that only one PC is useful for explaining variance within that sample set. PC1 versus PC2 scores plots were still constructed for SS3 for the purpose of comparison with the other sample sets. The explained variance plot also shows a distinct difference between the calibration and validation variances. This is indicative of the presence of outliers within the data set. Although initially thought to be due to the presence of pure milk powder in the data (due to these samples consistently grouping a large distance from the remaining data points), excluding these samples from the data had little improvement. This suggests that the points considered to be outliers were likely to be caused by the 'natural' variation which occurs between spectra of the same sample. This is because the only difference between the spectra of the same sample is the approximate $90^{o}$ rotation of the divet around a vertical axis which was performed between spectra. The large amount of 'natural' variation should be considered while interpreting the PCA scores plots of SS1, SS2 and SS3 as it becomes problematic when considering two samples with little difference in calcite concentration.

**Figure 4.6** contains three scores plots, each of which were created from spectra of sample set 1 (SS1). The purpose of this is to determine whether Raman spectroscopy coupled with principal component analysis could be a reliable and repeatable method for identification of spiked milk powder samples. The eight different colours on each scores plot represent a milk powder sample with a particular percentage of calcium carbonate within (w/w). The percentage of calcium carbonate contained within samples ranges from pure milk powder (i.e. 0% calcium carbonate) to 10% calcium carbonate. Even though it is unlikely that milk powder

will be supplemented with 10% calcium carbonate, this study is concerned with observing the ability of PCA for distinguishing between spiked samples, which requires a range of concentration differences. In this case, the differences include 0.5%, 2.5% and 5%. The results of PCA in **figure 4.6** show that the 10%, 5% and 2.5% samples are easily distinguished from one another. However, when the concentration difference between samples becomes 0.5%, distinguishing each sample becomes much more difficult. The 0.5%, 1%, 1.5%, 2% and 2.5% samples are grouped closely together. A gradient can be observed for two of the three plots with respect to PC1 for repetition 1 where the 0.5%, 1%, 1.5%, 2% and 2.5% samples are distributed in that order from left to right. Repetition 2 clearly shows disordered grouping with the 0.5% sample situated in the centre of the group with 2% and 1% samples located on the far left and 2.5% sample located on the far right. This scores plot also shows the presence of a distinct outlier from the 2.5% sample. The scores plot for repetition 3 should also be noted because the placement of the 5% sample relative to PC2 is significantly different from the other scores plots. Despite this apparent flaw and difficulty with distinguishing between samples with a difference of calcium carbonate content of only 0.5%, this technique still manages to effectively distinguish between pure milk powder and samples containing any amount of calcium carbonate. Otherwise, this scores plot similarly shows the 10%, 5% and pure milk powder samples form distinct clusters while the remaining samples are much more difficult to distinguish. The amount of variance explained by the first two PCs can be seen in the scores plots obtained from PCA. This value provides an indication of how well the variance within the data set is described by PCA. The scores plots indicate that the first two principal components in repetition 1 account for 78% of variance, in repetition two they account for 82% and for repetition 3 they account for 80%. These results suggest that the variance within these sets of data is explained moderately well.

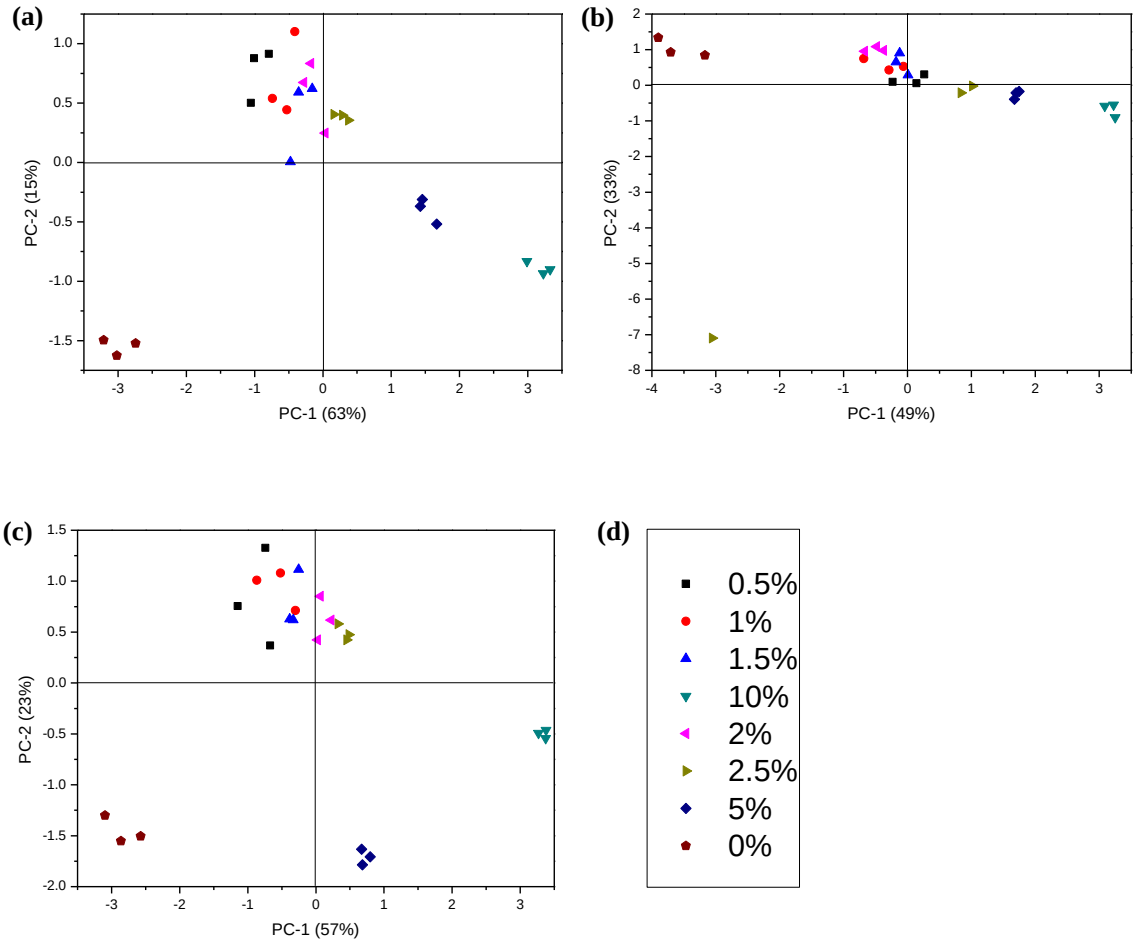



**Figure 4.6 Three scores plots produced from the spectra of sample set 1 (SS1): (a) repetition 1, (b) repetition 2 including the apparent outlier, (c) repetition 3 and (d) a key indicating the calcium carbonate concentration of each sample. Each plot was produced from 3 spectra taken from eight different spiked milk powder samples (including pure milk powder, 0%). This process was performed three times (the three repetitions). Each repetition was performed on a different occasion.**

**Figure 4.7** contains three scores plots of three repetitions collected using SS2. Similarly to the plots of SS1, these plots also contain very distinct 10% and 5% calcium carbonate groups as well as a large group of the remaining spiked milk powder samples. In general, these three plots all contain poorer distinction between samples with a difference of 0.5% calcium carbonate than the scores plots of SS1. The gradient which was present from 0.5% to 2.5% in repetitions 1 and 3 in SS1 is much less apparent in SS2 samples. The explained variance of the first two PCs is 83% for repetition 1, 72% for repetition 2 and 63% for repetition 3. Once

again, the pure milk powder is clearly distinguishable from any sample containing calcium carbonate.

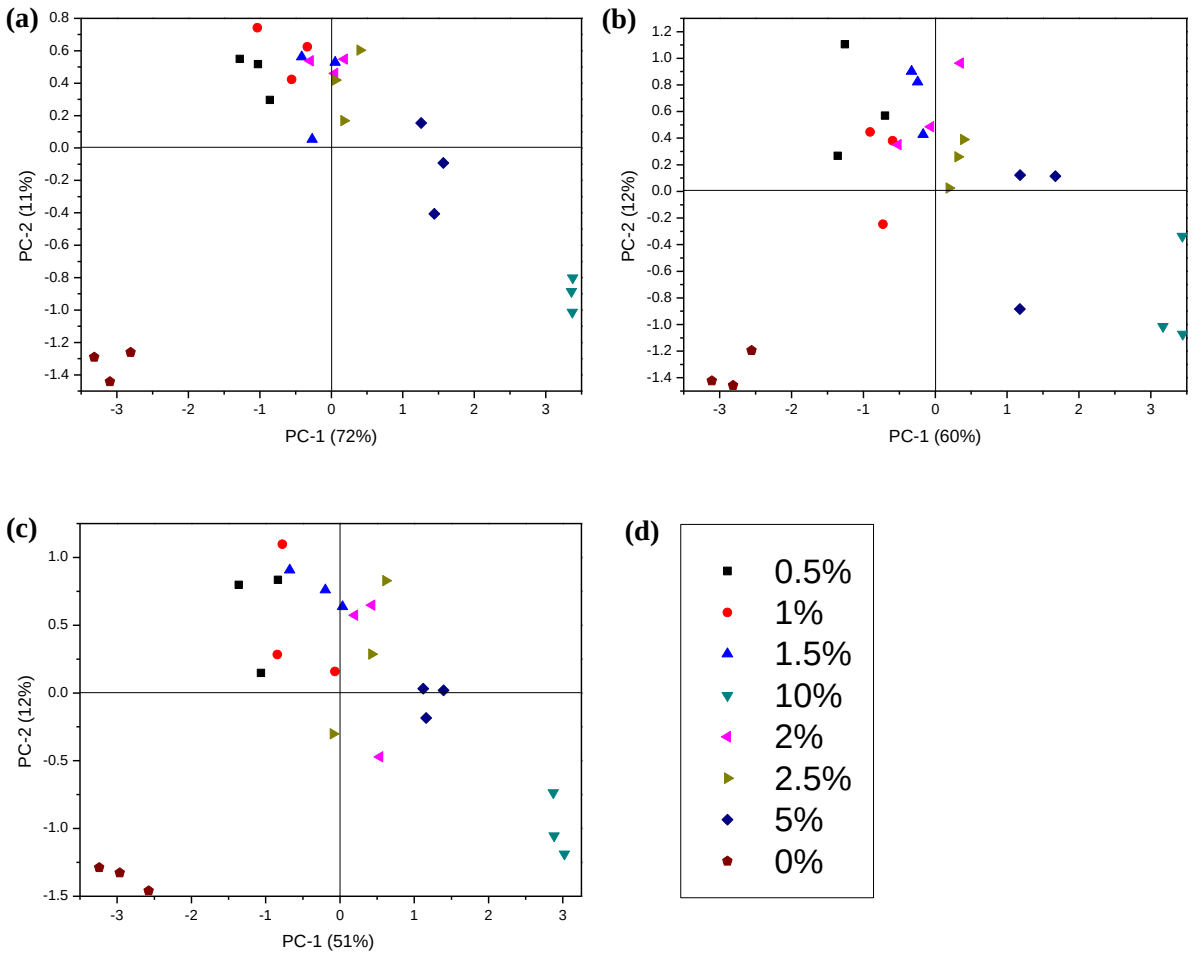



**Figure 4.7 Three scores plots produced from the spectra of sample set 2 (SS2): (a) repetition 1, (b) repetition 2, (c) repetition 3 and (d) a key indicating the calcium carbonate concentration of each sample. Each plot was produced from 3 spectra taken from eight different spiked milk powder samples (including pure milk powder, 0%). This process was performed three times (the three repetitions). Each repetition was performed on a different occasion.**

**Figure 4.8** contains the scores plots obtained from SS3. Like both SS1 and SS2, these scores plots also show the 5% and 10% calcium carbonate samples forming distinct groups. Pure milk powder clusters are once again distinct from all spiked samples. Repetition two is the only plot to show a calcium carbonate percentage gradient. Not only does it show a gradient with respect to PC1 but, with the exception of 1.5% sample points, also contains a notable

gradient with PC2. Both repetitions 1 and 3, however, show considerable overlap of sample groups in the range of 0.5 to 2.5% which obscures any notable gradient. PC1 and 2 for each of these plots explains comparatively little of the variance relative to SS1 and SS2. The first two PCs of repetition 1 explain 64% of the variance, 63% for repetition 2 and 59% for repetition 3; a relatively poor explanation of variance compared to the other sample sets.

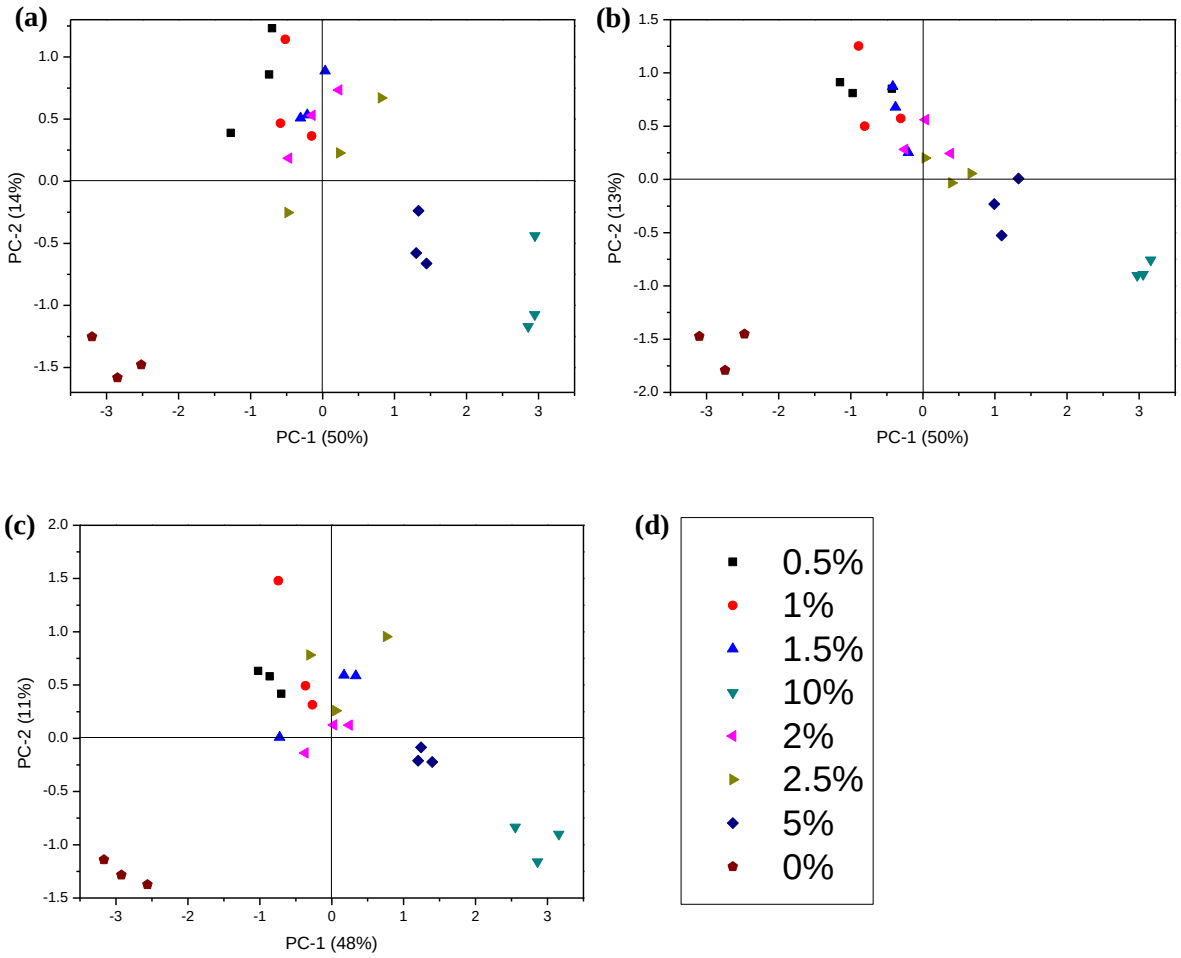



**Figure 4.8 Three scores plots produced from the spectra of sample set 3 (SS3): (a) repetition 1, (b) repetition 2, (c) repetition 3 and (d) a key indicating the calcium carbonate concentration of each sample. Each plot was produced from 3 spectra taken from eight different spiked milk powder samples (including pure milk powder, 0%). This process was performed three times (the three repetitions). Each repetition was performed on a different occasion.**

Results from principal component analysis of the three sample sets indicate that the 10%, 5% and pure milk powder samples form distinct clusters. The samples which have smaller

differences in calcium carbonate content are much harder to distinguish from one another due to overlap of sample clusters. PCs 1 and 2 explain the majority of variance in all sample sets and repetitions. The explained variance plots (only one provided) show a large decrease in the explained validation variance. This is possibly due to the presence of outliers within the data. The results suggest that spiked samples can be distinguished from milk powder and each other. However, pure milk powder data may need to be removed from the data matrix used for partial least squares analysis to allow for accurate quantification of calcium carbonate.

## 4.4 Univariate Analysis

### 4.4.1 Introduction to the technique

Univariate analysis is a technique used to analyse single variables. This involves plotting a univariate regression, where one variable is measured (the independent variable) and one response property is measured. In this thesis, univariate regressions are constructed using the calcium carbonate concentration as the independent variable and the integral of the peak at 1085 $cm^{-1}$ as the response property. The integral variable was chosen to measure the response to the changing concentration because it is the most prominent peak within the spectrum of calcium carbonate. It is also used because both principal component analysis (PCA) and partial least squares (PLS) analysis indicate that this peak is responsible for the majority of variance within spectra with changing calcium carbonate concentration. One of the main assumptions of this technique is that only one compound is contributing to the band used to create the regression. However, **figure 4.1** shows that both milk powder and calcium carbonate do contain bands at the 1085 $cm^{-1}$ region. Therefore this must be considered when interpreting results.

### 4.4.2 Results and discussion

A univariate regression was created for each sample set and the corresponding repetitions; however, not all will be presented in this thesis. However, the slopes, $R^2$ values and standard errors obtained from the 1085 $cm^{-1}$ integral versus concentration plots for all repetitions will be presented in this section. **Figure 4.9** contains the regression obtained from SS1 repetition 1. This plot indicates that there is a linear relationship between the integral values from the 1085 $cm^{-1}$ peak with the change in calcium carbonate concentration. **Table 4.1** summarises the

slope, $R^2$, percentage error of the slope and standard error values obtained from each of the sample set repetitions. The slope, standard error and $R^2$ values all vary greatly between sample set repetitions, with $R^2$ values often smaller than 0.80. The percentage error of the slope indicates that there is an uncertainty of approximately 13% the value of the slope in many samples. Limits of detection (LOD) and quantification (LOQ) were calculated using the standard error of the regression. LOD was considered to be three times the average standard error (**equation 4.2**) value plus the y-intercept value. LOQ was considered to be ten times the average standard error value plus the y-intercept. This gives LOD and LOQ values with the same units as the integrals (the dependent variable). The average slope, calculated from the gradients of all sample set plots (0.0051), was then used to convert the LOD and LOQ into values with units of: % (w/w) calcium carbonate. LOD and LOQ values were found to be 3.8% (w/w) calcium carbonate and 13% (w/w) calcium carbonate, respectively. The main assumption of calculating LOD and LOQ in this way is that the results obtained from each of the sample set repetitions are treated as though they are true repetitions (i.e. that each repetition was collected on the same day and each repetition came from the same sample set).

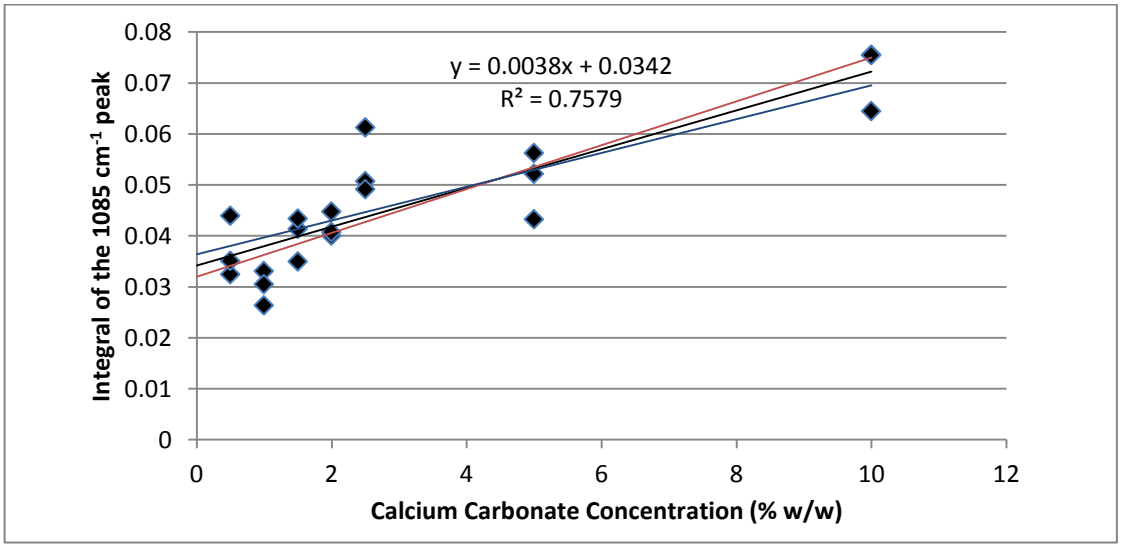



**Figure 4.9 Univariate regression of the values obtained from integration of the 1085 $cm^{-1}$ peak versus the concentration of calcium carbonate. The red and blue lines illustrate the standard error of the actual slope, which is considered to be the black line. Slopes for univariate regressions and the respective errors are listed in table 4.1.**

**Table 4.1 This table contains the slopes, percentage errors of slopes, $R^2$ values and the standard errors obtained from the univariate regressions of each sample set and the respective repetitions.**

| Sample Set | Repetition | Slope | Percentage error* | R-Squared | Standard Error** |
|---|---|---|---|---|---|
| SS1 | Rep 1 | 0.0038 | 13% | 0.76 | 0.0070 |
| | Rep 2 | 0.0068 | 13% | 0.77 | 0.012 |
| | Rep 3 | 0.0082 | 5.1% | 0.95 | 0.0060 |
| SS2 | Rep 1 | 0.0089 | 6.0% | 0.94 | 0.0075 |
| | Rep 2 | 0.0033 | 13% | 0.77 | 0.0059 |
| | Rep 3 | 0.0050 | 7.1% | 0.91 | 0.0050 |
| SS3 | Rep 1 | 0.0030 | 12% | 0.78 | 0.0051 |
| | Rep 2 | 0.0043 | 7.6% | 0.90 | 0.0046 |
| | Rep 3 | 0.0030 | 13% | 0.76 | 0.0055 |
| Average | N/A | 0.0051 | 10% | 0.84 | 0.0065 |

*Refers to the percentage error of the slope and was calculated from the standard error of the slope.

**Refers to the standard error of the regression and was used to calculate the LOD and LOQ. These values were **not** used to calculate the percentage error of the slope.

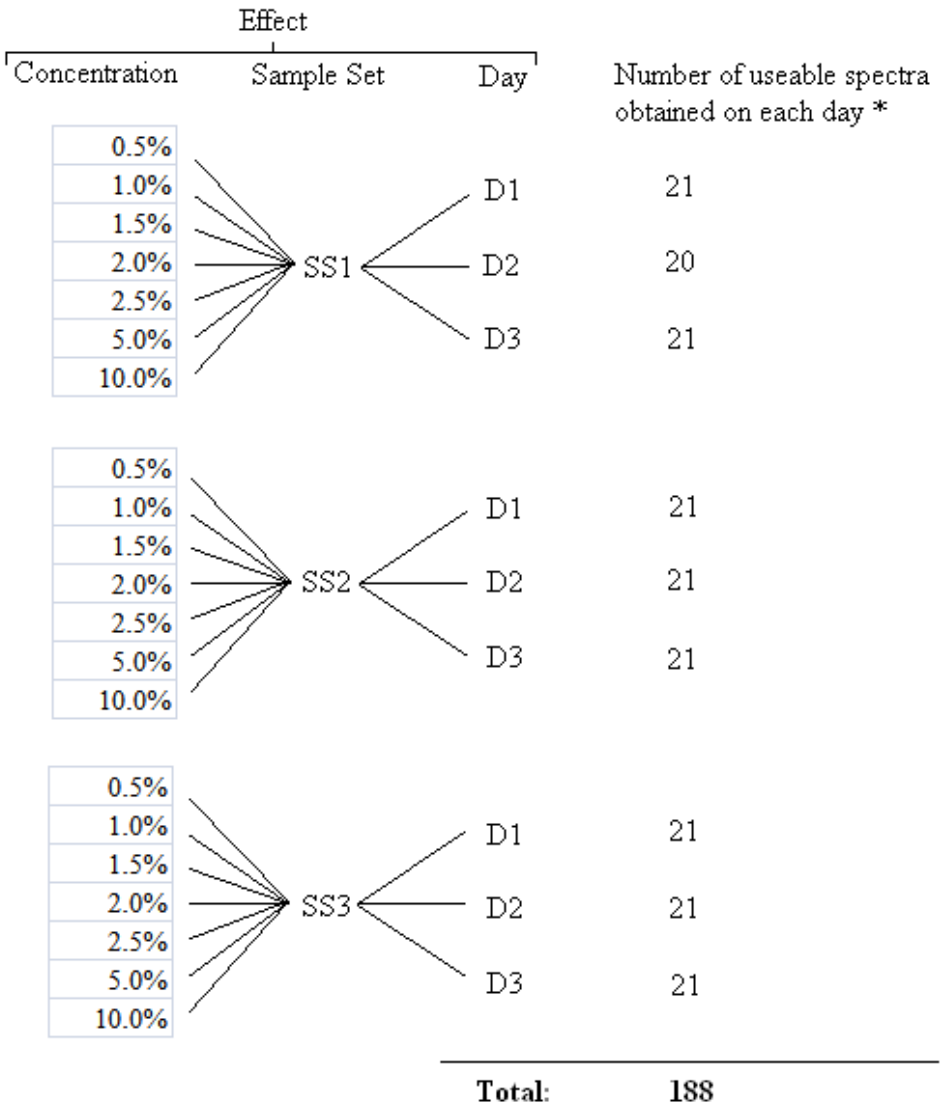
$$Standard\ Error = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n-2}}$$

**Eq. 4.2**

### 4.4.3 Analysis of variance (ANOVA)

The analysis of variance (ANOVA) was performed on the univariate analysis data in order to determine the possible causes of the large amount of variation in the integral data. Prior to the ANOVA test, however, the data required transformation using the natural logarithm function. This is because the raw integral data was skewed, with the majority of samples with 1085 $cm^{-1}$ peak integral values lower than the median. Transforming the data produced a normal distribution within the data set which allowed an ANOVA test to be performed. ANOVA is a statistical process which is used to analyse the variations in a particular response and relates it to several effects (variables) which may be associated with the variation.[48] The influence of an effect on the variation is gauged by its assigned level of significance. The significance is determined from the p-value of the respective effect. In order to conclude that an effect is

significant the associated p-value must be small. The process of finding this value begins with defining the sources of variation (the effects) and the degrees of freedom (DF). The degrees of freedom denote the number of ways in which a particular effect can influence variation in the dataset. For each effect this is considered to be equal to the number of individual units present minus one. For example, there are 7 concentrations; therefore there are 6 degrees of freedom for this effect. The sum of squares (SSq) is then calculated, which illustrates the amount of response variation associated with a particular effect. The response variance is then calculated by dividing the sum of squares by the associated number of degrees of freedom for that particular effect.[48] This ratio is referred to as the mean square (MS). F-ratios are then calculated by dividing MS(effect) by the relative error term. These ratios are useful because they compare the amount of variance caused by a particular effect with the residual variance; hence, it provides information on the significance of the variance caused by a particular effect. The p-value is then calculated and is an estimation of the probability of obtaining the F-ratio value if the null hypothesis is true. In this study, the response is the integral of the 1085 $cm^{-1}$ peak. There are 3 different 'sample sets', each of which consist of 7 concentrations of calcium carbonate: 0.5, 1, 1.5, 2, 2.5, 5 and 10% (w/w). Each sample set is analysed on 3 different occasions; for simplicity these are referred to as days 1, 2 and 3 because a large majority of 'occasions' took place on different days. They are also referred to as days rather than 'repetitions' during the ANOVA because it highlights the variable of interest. It is important to note that the concentration levels and sample set preparation methods are fixed and the days were chosen randomly. **Figure 4.10** illustrates the main effects present in this model; calcium carbonate concentration, sample set and day.




* Three spectra were obtained for each of the 7 concentrations on each day. SS1 Day 2 contained a spectrum which was considered an outlier and unusable for univariate analysis, PLS analysis and ANOVA tests.

**Figure 4.10 Illustrates three possible effects of the ANOVA test: concentration, sample set and day. The sample sets are comprised of seven samples each containing a different concentration (w/w) of calcium carbonate. Each sample set is also analysed on three different occasions, referred to as days 1, 2 and 3 (D1, D2 and D3). On each day, three spectra are collected for each concentration, hence a total of 21 spectra were obtained on all days. SS1 D2 had only 20 usable spectra because one of the 2.5% spectra was a significant outlier.**

**Table 4.2 Results of an ANOVA test. The response variation is considered to be the variation in the integral of the 1085cm$^{-1}$ peak. The effects are considered to be the concentration (C), the sample set (S), the day the spectrum was taken from the corresponding sample set (D(S)), the interaction between the concentration and sample set (C*S) and the interaction between the concentration and the day (C*D(S)). Interaction refers to the combined effect of the two variables on the value of the integral. The abbreviations refer to degrees of freedom (DF), unadjusted sum of squares (SSq), adjusted sum of squares (Adj SSq) and the adjusted mean square (Adj MS). Adjusted SSq values were calculated using the natural logarithm of the integrals due to the skewed integral data, as discussed earlier in this section. The P-values indicate that the concentration, day and the interaction between the two effects are significant.**

| Source | DF | SSq | Adj SSq | Adj MS | F-ratio | P-value |
|---|---|---|---|---|---|---|
| C | 6 | 15.76293 | 15.71399 | 2.61900 | 42.84 | 0.000* |
| S | 2 | 11.21020 | 11.49050 | 5.74525 | 2.62 | 0.152* |
| D(S) | 6 | 13.10214 | 13.15092 | 2.19182 | 35.86 | 0.000* |
| C*S | 12 | 0.45052 | 0.45052 | 0.03754 | 0.61 | 0.816* |
| C*D(S) | 36 | 2.20282 | 2.20282 | 0.06119 | 3.20 | 0.000 |
| Error | 125 | 2.39049 | 2.39049 | 0.01912 | | |
| Total | 187 | 45.11888 | | | | |

*Not an exact F-test as an adjusted sum of squares was used to calculate the mean square. The adjusted sum of squares was calculated using the natural logarithm of the integral rather than the value of the integral itself.

The ANOVA indicates that both the concentration of calcium carbonate and the day have a significant effect on the integral of the 1085cm$^{-1}$ peak, while the sample set does not (**table 4.2**). It also indicates that there is significant interaction between concentration and day which is affecting the integral values. This was investigated through use of an interaction plot, presented in **figure 4.11**. The plot shows the median natural logarithm of the integral for each concentration on each day. There appears to be very little change in the average integral obtained between days 1 and 2, with the exception of the 2.5% samples. However, there is a consistent decrease in the average integral value between days 2 and 3. **Table 4.3** illustrates the significance of the changes in integral values with each day. This table presents the 95% confidence intervals of the ratio between integrals measured on different days. For example, for the 0.5% concentration level for $y_{D2/D3}$ one is 95% confident that the true median of y (the integral of the 1085 cm$^{-1}$ peak) on day 2 is between 1.39 times and 1.79 times the median on

day 3. This table indicates that the change in integral values from day 1 to day 2 is not significant except for the 2.5% samples. It also indicates that there is a significant decrease in the integral values from day 2 to day 3, with the exception of the 10% samples.

These results indicate that another study is likely to be required to determine whether the consistent decrease in integral values from day 2 to day 3 is due to the influence of the performance of the instrument; the age of the sample; or another cause. One of the reasons for this is due to the use of too few day 3 'repetitions' (i.e. SS1 Day 3, SS2 Day 3 and SS3 Day 3). More repetitions may have given integral values which matched day 1 and day 2 values more closely, thus indicating that the decrease was due to variation in the instrument performance. Age may be the cause of the decreasing integral values because each day corresponds to a different relative age of the sample. Sample set 1, for example, was analysed on three different days (days 1, 2, and 3). Day 3 spectra were always obtained after day 2 spectra and day 2 spectra were always obtained after day 1 spectra. Therefore, the increase in day number also corresponds to an increase in the age of the sample. Whether this had an effect on the peak signal could be assessed by increasing the number of days on which the samples were analysed and uniformly distributing the occasions when analysis occurs. Data obtained in this thesis are not sufficient to come to any conclusions about the cause of the consistent decrease in integral value of day 3 spectra because the days on which spectra were collected were chosen randomly and the samples would have to be analysed on several more occasions.

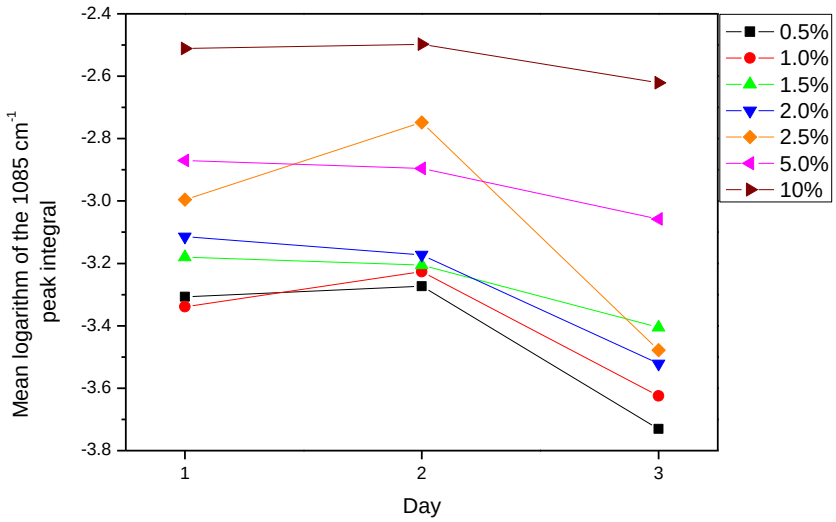



**Figure 4.11 Interaction plot of the average logarithm of the 1085 $cm^{-1}$ peak integral versus the day on which the corresponding spectra were obtained. Spectra have been divided into groups depending on the concentration of calcium carbonate (% w/w) in the analysed sample. The key denotes the different concentration groups.**

**Table 4.3 Illustrates the changes in the 1085 $cm^{-1}$ peak integral between days 1 and 2 and between days 2 and 3. This is achieved by calculating the 95% confidence interval for D1/D2 and D2/D3 ratios. For example, $y_{D1/D2}$ refers to the ratio of the median 1085 $cm^{-1}$ peak integrals obtained on day 1 with those obtained on day 2. Therefore $y_{D1/D2}$ 95% CI refers to the 95% confidence interval for this ratio, which is presented as the lower and upper limits of the confidence interval.**

| Concentration* | $y_{D1/D2}$ 95% CI | $y_{D2/D3}$ 95% CI |
|---|---|---|
| 0.50 | 0.851, 1.10 | 1.39, 1.79 |
| 1.0 | 0.787, 1.02 | 1.31, 1.69 |
| 1.5 | 0.903, 1.17 | 1.07, 1.39 |
| 2.0 | 0.932, 1.20 | 1.25, 1.61 |
| 2.5 | 0.684, 0.890 | 1.82, 2.37 |
| 5.0 | 0.903, 1.17 | 1.04, 1.34 |
| 10 | 0.868, 1.12 | 0.996, 1.29 |

* Units: % (w/w)

## 4.5 Partial Least Squares (PLS) Analysis

### 4.5.1 Introduction to the technique

Partial least squares is used to calculate new variables, known as factors, which maximise covariance between two matrices.[9] Factors may also be referred to as latent variables or PLS-components and were calculated using the NIPALS algorithm in The Unscrambler X. It is important to note that they are fundamentally different from principal components calculated during PCA due to the way in which they are calculated, therefore PLS factors and principal components are not interchangeable terms.[48] PLS essentially finds the relationship between two matrices which are referred to as matrices X and Y. Matrix X, in this case, refers to a matrix of spectroscopic data collected from several samples containing different concentrations of calcium carbonate. Matrix Y refers to a simple matrix containing only an array of calcium carbonate concentrations. A useful analogy for understanding PLS and how it differs from PCA, used by Esbensen, suggests initially viewing the technique as two simultaneous PCA analyses, PCA of X and PCA of Y.[48] In doing so, one can understand how the X-matrix scores (T) and loadings (P) arise, as well as the Y-matrix scores (U) and loadings (Q). The X-matrix also has 'W', an alternative loading matrix in addition to the P-loading, referred to as the loading weight. **Figure 4.12** contains a schematic overview of PLS relating each of these terms. Now that T, P, W, U and Q have been introduced, it is important to emphasise that viewing PLS as two simultaneous PCA analyses is not entirely correct. Decomposition of X and Y matrices is actually performed in an interdependent manner. PLS really uses the u-score vectors ($u_1$) instead of the t-score vectors ($t_1$) in the decomposition of the X-matrix. Therefore the Y-data structure directly guides the decomposition of X. $U_1$ is then replaced by $t_1$ at a particular stage of the algorithm when the Y matrix is decomposed. In summary, $u_1$ is initially used to influence the X-decomposition to find the X-loadings 'w'. T-vectors are then calculated using the w-vector. The t-vector is then used instead of $u_1$ for the decomposition of the Y matrix. Following this process, a final set of t, w, u and q vectors are calculated for the relevant PLS factor (*f*). Further information about PLS algorithms can be found in reference 48.

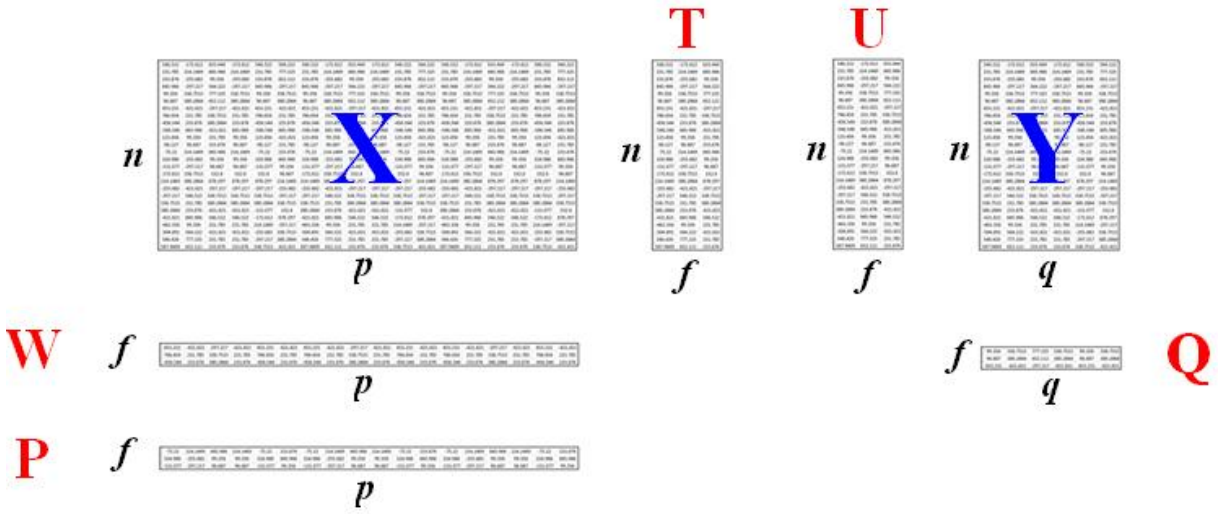



**Figure 4.12 A schematic summary of PLS data decomposition based on the matrix created from two matrices, X and Y. In this thesis, matrix X is constructed using the collected spectroscopic data and matrix Y is constructed from the concentration variable data. P and T are the loading and scores matrices from X, respectively. W is the loading weight matrix for X. Q and U are the loading and scores matrices of Y, respectively. The number of samples is represented by *n*. The number of spectroscopic variables (Raman intensities at differing Raman shifts) is represented by *p*. The number of concentration variables is represented by *q*. The optimum number of PLS factors is represented by *f*. This picture is based on images from Esbensen[48] and McGoverin.[9]**

PLS analyses may be explored in several ways including via the scores, loadings, explained variance (scree plot), X-Y relationship outliers and predicted versus measured plots. The scores plot involves plotting two factors against each other, similar to the scores plots calculated in PCA. The explained variance plot illustrates the amount of variance each factor explains. PLS loadings convey the relationship between the raw spectroscopic data and its scores and may be interpreted in a similar fashion as in PCA.[48] Therefore the structure within the spectroscopic data matrix correlated to the dominant variation with changing calcium carbonate content will have the largest value coefficients in the loadings plot. Regression coefficients in this plot describe vectors which are used to calculate concentrations which are presented as predicted concentrations in the predicted versus measured plot. The predicted versus measured plot expresses the relationship between predicted and measured concentrations. Outliers within the model can be identified using the X-Y relationship outliers plot. The Unscrambler X (CAMO Technologies Inc., Woodbridge, USA) describes this plot as

a summary of the regression relationship between X and Y matrices along a particular factor of the PLS model. It involves plotting u-scores versus t-scores. U-scores summarise the part of the structure in Y which is explained by X along a given factor. T-scores capture the part of the structure in X which is most predictive for Y. It is primarily used to identify outliers which have been used to construct the X-Y relationship. Calibration error can be parameterised by the root mean square error of calibration (RMSEC), the square root of the average of the sum of squares differences between the predicted ($\hat{y}_i$) and the reference concentrations ($y_i$), as illustrated by **equation 4.3**.[9] This error has the same units as the reference concentration values. Root mean square error of prediction (RMSEP) and root mean square of cross validation (RMSECV) are analogous parameters calculated from prediction and cross validation results, respectively. When performing a PLS analysis, the appropriate number of factors should be used for modelling the data in order to avoid overfitting or underfitting the data. Overfitting occurs when too many factors are used for calibration and noise is modelled. Underfitting occurs when too few factors are used and incomplete modelling of the data structure takes place. The appropriate number can be determined from the plots of RMSE versus the number of factors used. However, this must be performed using the RMSE of validation (i.e. RMSEP and RMSECV). This is because the error of calibration, RMSEC, will continue to decrease with the increasing number of factors used until the RMSEC reaches a value of zero. However, the validation errors, RMSEP and RMSECV, will only decrease until noise within the calibration set begins to be modelled.[9] Hence the major inflection point within the plots of RMSEP or RMSECV versus the number of factors can be used to determine the appropriate number of factor to use.[9, 29] The purpose of this project is to determine whether Raman spectroscopy and PLS analysis can be used to quantify the calcium carbonate concentration in a spiked milk powder sample.

$$RMSEC = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}}$$

**Eq. 4.3**

### 4.5.2 Results and discussion

Each of the mixtures and their respective replicates (sample sets 1-3, repetitions 1-3) were processed using partial least squares analysis. All spectra used in PCA were also used for PLS except for one 2.5% calcium carbonate mixture from sample set 1 (SS1) repetition 2, which was considered to be an outlier after inspection of the PCA scores plot and visual inspection of the spectrum. PLS scores, loadings, explained variance, X-Y relationship outliers, RMSE versus number of factors and predicted versus measured plots were analysed. Results were validated using both leverage corrected validation and full cross validation. This thesis will only present the plots obtained from analysis of SS1 repetition 1; however, the slope, RMSEC, RMSEP, RMSECV and $R^2$ results obtained from the predicted versus measured concentration plots of each sample set and its respective repetitions will be presented.

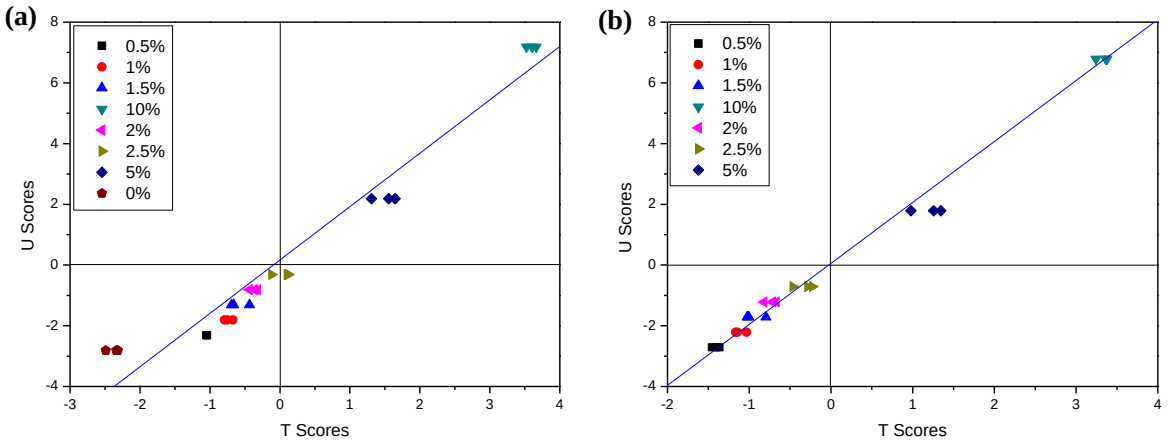



**Figure 4.13 (a) X-Y relationship outliers plot constructed with pure milk powder (0%) included. (b) X-Y relationship outliers plot constructed with pure milk powder data excluded. The key in each plot indicates the calcium carbonate concentration (% w/w) of each sample.**

**Figure 4.13** contains two X-Y relationship outliers plots. These plots summarise the regression relationship between predicted and measured concentrations and can be used to identify outliers which are present in a model. Initially, the spectroscopic data from pure milk powder was included in the matrix used for PLS analysis. However, **figure 4.13a** indicates that the pure milk powder samples behave like an influential outlier which adjusts the component orientation. Therefore pure milk powder was removed from the model with intent to avoid this effect. **Figure 4.13b** shows an X-Y relationship outliers plot calculated from PLS

analysis without using the pure milk powder data. This plot now exhibits a linear relationship between t-scores and u-scores and the absence of any further outliers. Therefore pure milk powder was excluded from all PLS analysis of all sample sets.

The PLS scores plot of SS1 repetition 1 is shown in **figure 4.14a.** This plot shows that each of the spiked milk powder mixtures can be separated based on the calcium carbonate concentrations. Factor 1 is shown to account for 99% of the variance with factor 2 accounting for the remaining 1%. This indicates that a total of 100% of the variance is explained by only factors 1 and 2. This is supported by the explained variance plot in **figure 4.14b. Figure 4.15** shows a loadings plot of factor 1 which indicates that the Raman shift at 1085 $cm^{-1}$ has the largest number of regression coefficients. This suggests that the peak at 1085 $cm^{-1}$ has the greatest chemical significance when calibrating the predicted versus measured concentrations plot and hence is the most useful for quantifying calcium carbonate when using factor 1.

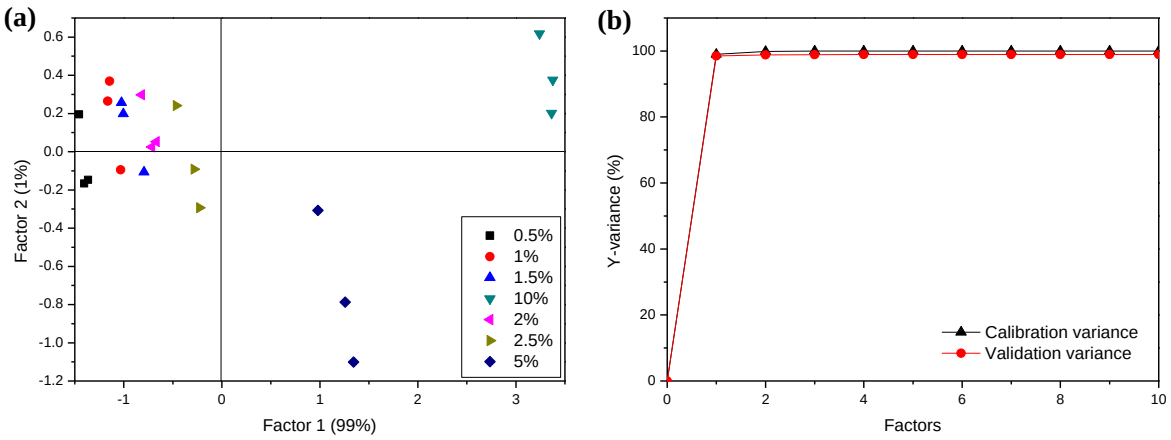



**Figure 4.14 (a) scores plot obtained from PLS analysis of SS1 repetition 1. The key indicates the calcium carbonate concentration (% w/w) of each mixture. (b) Explained variance plot obtained from PLS analysis of SS1 repetition 1. The calibration variance is in black and the validation variance in red.**

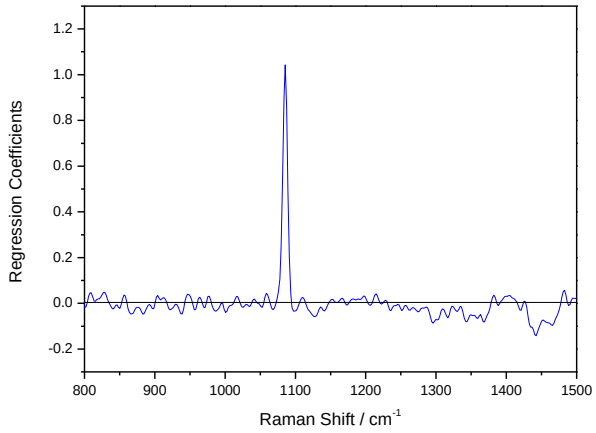



**Figure 4.15 A loadings plot constructed from factor 1 of SS1 repetition 1 after PLS analysis.**

The predicted versus measured plot illustrates the relationship between predicted and measured concentrations of calcium carbonate. The optimum number of factors to use to calibrate this plot was determined from the explained variance curve after cross validation and a plot of RMSE versus the changing number of factors (**figures 4.14b** and **4.16a**, respectively). The explained variance plot indicates that 99% of variance is explained by factor 1, and factor 2 explains the remaining variance. The RMSE versus number of factors plot (**figure 4.16a**) shows RMSEC (black) decreases with increasing numbers of factors until an error of 0% is reached; however, RMSECV only decreases until the noise becomes modelled. The appropriate number of factors can be considered to be the point at which there is no significant difference in RMSECV with increasing number of factors. The plot shows very little difference in RMSECV between factor 1 and the remaining factors. This occurred for each other sample set (not shown) and indicates that using one factor is the optimum number to use. **Figure 4.16b** shows the predicted versus measured plot of SS1 repetition 1. The slope of this plot is 0.99 with an R-square value of 0.99 which suggests that calcium carbonate content was able to be effectively predicted using PLS and the plotted line has a good fit. The slopes, $R^2$ values and the errors of all sample sets are summarised in **table 4.4**. The average slope obtained from all samples is 0.99. The errors of the predicted concentrations are illustrated by RMSEC, RMSEP and RMSECV values. The average RMSEC, RMSEP and RMSECV values obtained are 0.34, 0.38 and 0.52% (w/w), respectively. The errors are small and are in the same order of magnitude as the respective errors reported in a study of Raman spectroscopic quantification of fat and protein in milk

powder using PLS.[36] The average LOD and LOQ values for PLS are 1.0% and 3.4%, respectively. The values were calculated using the RMSE rather than standard error as explained in **section 4.4.2**. The main assumption in doing so is that there is little difference between RMSE and standard error values. This is justified by the small difference in the way in which these errors are calculated, exemplified by **equations 4.2** and **4.3**. These results suggest that it is also possible to predict calcium carbonate content using Raman spectroscopy and PLS.

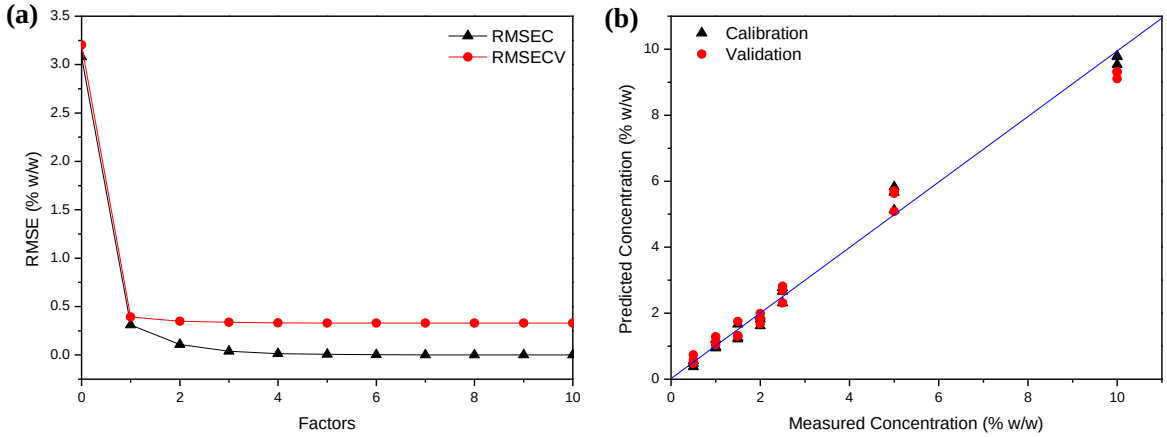



**Figure 4.16 (a) Plot of root mean square error of prediction versus the number of factors used for calibration of the predicted versus measured plot. (b) predicted versus measured plot calibrated using one factor.**

**Table 4.4 The slopes, RMSEC, RMSEP, RMSECV and R-square values obtained from the predicted versus measured plots of all samples. This table includes values from plots calibrated using one factor.**

| | Slope | RMSEC* | RMSEP* | RMSECV* | R-squared |
|---|---|---|---|---|---|
| **SS1 Rep 1** | 0.99 | 0.31 | 0.35 | 0.39 | 0.99 |
| **SS1 Rep 2** | 0.96 | 0.65 | 0.71 | 0.76 | 0.96 |
| **SS1 Rep 3** | 0.99 | 0.17 | 0.19 | 0.47 | 0.99 |
| **SS2 Rep 1** | 0.99 | 0.20 | 0.23 | 0.27 | 0.99 |
| **SS2 Rep 2** | 0.99 | 0.35 | 0.41 | 0.49 | 0.99 |
| **SS2 Rep 3** | 0.99 | 0.29 | 0.32 | 0.48 | 0.99 |
| **SS3 Rep 1** | 0.99 | 0.33 | 0.38 | 0.54 | 0.99 |
| **SS3 Rep 2** | 0.98 | 0.39 | 0.44 | 0.54 | 0.98 |
| **SS3 Rep 3** | 0.99 | 0.34 | 0.37 | 0.71 | 0.99 |
| **Mean** | 0.99 | 0.34 | 0.38 | 0.52 | 0.99 |
| * Units: % w/w | | | | | |

### 4.5.3 Comparison of PLS with univariate analysis

The results of univariate and PLS analysis suggest that both techniques have the ability to quantify calcium carbonate which has been mixed with milk powder. Comparison of $R^2$ and error values indicates that PLS is the more effective technique to use for this purpose. The LOD and LOQ values obtained for each technique illustrate this more clearly. The average LOD and LOQ obtained from univariate analysis are 3.8 and 13% (w/w) calcium carbonate, respectively. The average LOD and LOQ obtained for PLS analysis are 1.0 and 3.4% (w/w) calcium carbonate, respectively. As well as the lower LOD and LOQ, PLS analysis obtained more consistent gradients and $R^2$ values from the predicted versus measured concentration plots. This indicates that PLS is likely to be more effective than univariate analysis for determining calcium carbonate concentrations in spiked milk powder samples. However, written conversations with Steve Holroyd from Fonterra have confirmed that retail milk powder fortified using calcium carbonate contain less than 3.2% (w/w) calcium carbonate. Despite being capable of quantifying calcite in spiked milk powder samples, the LOQ of PLS analysis is 3.4%. This suggests that use of Raman spectroscopy and PLS would not be able to accurately quantify calcite in retail milk powder samples. RMSE values obtained in this investigation are comparable to those found in the literature during the quantification of fat and protein using Raman spectroscopy and PLS analysis.[36] This suggests that the calculated errors are accurate and supports the aforementioned conclusion.

## 4.6 Conclusions

Principal component analysis can be used to identify spiked milk powder samples as all samples with added calcium carbonate form clusters distinct from the pure milk powder cluster. Samples with larger differences in calcite content were also able to be distinguished from each other effectively. However, samples which differed in concentration by only 0.5% were difficult to separate by PCA. After excluding the spectroscopic data of pure milk powder, investigation of PLS and univariate analysis indicates that both of these techniques can be used with Raman spectroscopic data to quantify calcite within milk powder samples. PLS was found to be more effect for this purpose than univariate analysis after comparison of the LOD and LOQ values obtained using each technique. However, the LOQ of PLS is larger than the quantity of calcium carbonate typically used to fortify retail milk powder. This

suggests that both PLS and univariate analysis would not be useful for the quantification of the small concentrations of calcium carbonate used in milk powder samples prepared for retail.

# Chapter 5 Cheese Analysis

## 5.1 Introduction

Quality control of dairy products, such as cheese, is an important part of the manufacturing process. There are many types, including cheese-type foods that are a blend of natural cheese and other food ingredients. Analytical techniques such as near-infrared (NIR) and mid-infrared (MIR) spectroscopy have been previously used to study cheese contents such as protein and fat.[49, 50] By comparison, Raman spectroscopy has garnered relatively little attention in this field. The technique has, however, been used in the analysis of several food ingredients of which cheese-type foods may be comprised. The components of interest include: anhydrous milk fat (AMF),[32] rennet casein (RC),[51] sucrose (SUC) [52] and cornstarch (CS).[53] Raman spectroscopy has been used to investigate these components in dairy products.

AMF has previously been examined using Raman spectroscopy in a study trying to determine whether solid fat content (SFC) and triacyglycerol (TG) content in AMF could be predicted using the technique.[32] This investigation involved heating several samples of AMF and recording spectra at 5 $^{o}$C increments. This study was discussed in more detail in **section 1.4** therefore the results will be summarised briefly in this section. The result of this research suggested that SFC could be predicted from physically heterogeneous samples using Raman spectroscopy at room temperature. Quantification of TGs was found to be more difficult. It was found to be impossible for some TGs but possible for others. Quantification of TGs was found to be most effective when samples were liquid, hence at higher temperatures.

Casein has been previously studied using Raman spectroscopy in order to investigate its secondary structures.[51] Caseins constitute approximately 80% of bovine milk proteins; two major components of which are $\alpha_{S1}$ and β-casein. The conformations of casein molecules are poorly understood, but are known to differ in primary and secondary structures. Based upon sequence studies, the secondary structure of $\alpha_{S1}$ at pH 6.8 is estimated to consist of 8-17% β-structure and 13-22% helix structure. Secondary structure of β-casein at pH 6.8 is estimated to consist of 0-13% β-structure and 1-20% helix structure.[51] The 1580 – 1720 cm$^{-1}$ region of the spectrum was of particular interest. Spectra received resolution enhancement of amide I bands

via Fourier deconvolution, a method used for general smoothing or removal of specific frequency components.[29] The results of the investigation using Raman spectroscopy suggested that $\alpha_{S1}$ and β casein have around 10% helical structure, 20% β-structure and from 20-35% turn structures.

Sucrose content within 'lemon-type' soft drinks has been quantified using Raman spectroscopy.[52] The purpose of this is to find an efficient method for quality assurance. An 830 nm laser wavelength was use to excite the sample and a CCD detector was used to collect dispersed light. Quantification was achieved by preparing a calibration curve using a series of sugar solutions ranging from 0 and 15.0 g/100 mL and use of partial least squares analysis (PLS) to correlate Raman spectra with sugar content. The result of this investigation suggested that the predicted sucrose content lay within 5.5% of that claimed by the manufacturers.

Cornstarch and cassava starch were analysed using FT Raman spectroscopy to investigate whether the two starches could be separated based on amylose content and whether amylose content could be quantified.[53] The experimental procedure involved identifying the quantity of amylose in the sample through the use of a Shimadzu model UVPC 1601 spectrophotometer using the working curve method, then used a Bruker RFS 100 FT-Raman spectrometer with PCA to observe whether or not FT-Raman spectroscopy together with chemometric analysis could be used to achieve accurate quantification. Chemometric analysis involved using PCA and PLS analysis to separate the samples and quantify the amylose content, respectively. The results of this experiment indicate that cornstarch, cassava starch and mixtures of both can be separated effectively using PCA. Quantification of amylose was achieved through use of Raman spectroscopy and PLS analysis. The calculation of the root mean square error of calibration (RMSEC), root mean square error of prediction (RMSEP) and relative errors indicated that the errors of this technique are small and are similar to those of colorimetry (the standard method). Therefore, this investigation suggests that Raman spectroscopy and chemometrics can be used for separation of starch samples and quantification of amylose. The advantage of using FT Raman is the lack of the requirement for pre-treatment, hence it gives faster analysis and it is non-destructive.

This chapter investigates whether bi-component cheese systems can be distinguished based on component concentrations using Raman spectroscopy with principal component analysis, and investigates whether chemical or physical interactions between components can be observed in both bi and tri-component systems. Creating model systems such as these gives the ability to easily identify flaws in technique due to the limitation of variables affecting results. All mixtures contain AMF and either one or two other components. **Component B** is used to generally refer to CS, RC or SUC.

## 5.2 Differentiation of Cheese Mixtures

Spectra were obtained from binary mixture set A (BMA) using FT-Raman spectroscopy. The purpose of this investigation was to observe whether the different AMF and component B mixtures could be reliably distinguished by component concentrations. PCA (described in **section 4.3.1**) enables this to be done by breaking down the variance between spectra into principal components. The scores plot separates the mixtures on the basis of amount of variance with respect to principal components Principal component 1 (PC1) and principal component 2 (PC2) usually explain the majority of the variance within a dataset; therefore by plotting PC1 versus PC2 (a scores plot), any significant variations between spectra could be identified. The loadings plot illustrates the regions of the spectra which exhibit the largest amount of variance. The influence plot illustrates whether or not there are any residual factors affecting variance other than the principal components. This is supported by the explained variance plot, which shows the amount of variance which is explained by each principal component.

PCA of AMF/CS mixtures (**figure 5.1**) gives a scores plot illustrating an almost linear distribution of points with respect to both PC1 and PC2 (with the exception of pure AMF). All mixtures are easily distinguished from one another and a consistent increase in the AMF proportion within the mixture can be observed from left to right along the x-axis (PC1). This indicates that separation of mixtures based on component concentrations can be achieved. PC2 describes the second-most variance between mixtures yet it does not produce a discernible pattern relating the mixture grouping with the increasing proportion of AMF. However, it does allow for a greater ability to distinguish between mixtures. The loadings plot relates the

Raman shift with amount of variance. In the AMF/CS mixtures it is apparent that the largest amount of variance occurs at two regions: ~2900 $cm^{-1}$ and ~477 $cm^{-1}$. Incidentally, it is the 477 $cm^{-1}$ region which is used as a characteristic peak to indicate the presence of CS in both 2D and 3D contour maps, discussed in Chapter 6. The influence plot lacks any significant pattern which suggests that there are no residual factors affecting variance within the sample. As well as the scores plot, the plot of explained variance indicates the amount of variance which is explained by each PC. From this plot it can be seen that PCs 1 (89%) and 2 (7%) explain 96% of the total variance, which suggests that these two PCs describe the variance within the data better than any other combination of PCs.

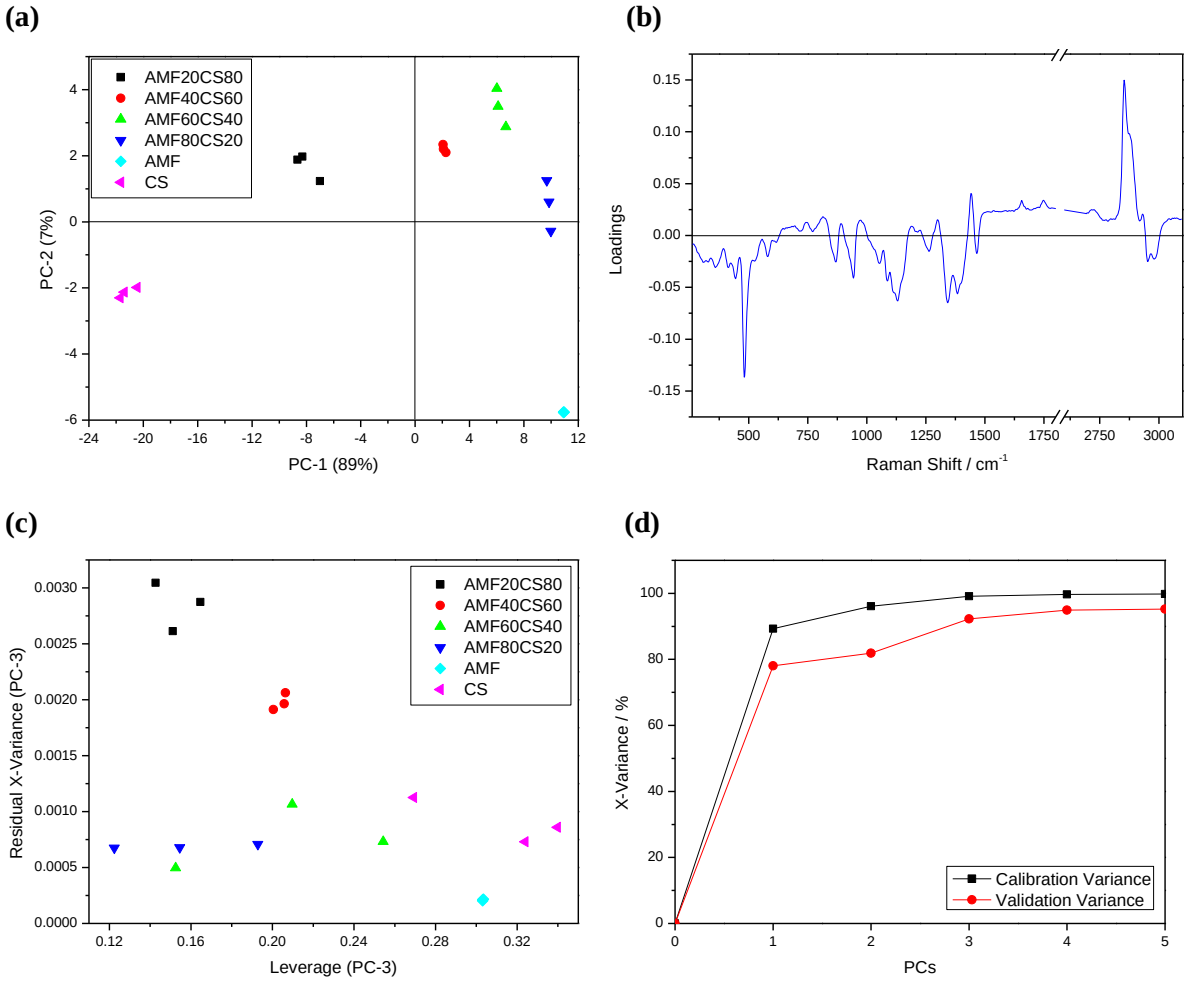



**Figure 5.1 Results of principal component analysis for AMF/cornstarch. This figure contains: (a) the scores plot, (b) the loadings plot, (c) the influence plot, and (d) the explained variance plot. The keys in the scores and influence plots indicate the symbols which correspond to each sample.**

The scores plot of AMF/RC (**figure 5.2**) mixtures also shows the mixtures as being very distinct from one another. However, unlike the AMF/CS mixtures, there is no pattern relating sample grouping with AMF concentration along the x-axis (PC1) or the y-axis (PC2). This is due to the AMF40RC60 mixture grouping between pure RC and AMF20RC80 along the x-axis. The difference between AMF40RC60 and AMF20RC80 along this x-axis is very slight. However, these two mixtures are very distinct with respect to PC2, which allows separation of the clusters to be observed. The explained variance and scores plot both indicate that PC1 explains much less variance (and PC2 explains more) than in the example of AMF/CS mixtures. In AMF/RC mixtures PC1 explains 78% of the total variance and PC2 explains 20%, giving a total of 98% between them. The loadings plot indicates that much of the variance occurs at ~2900 $cm^{-1}$ with much of the remainder spread between several broad peaks as well as a sharp peak at ~1003 $cm^{-1}$. The influence plot for these mixtures indicates the presence of an outlier within the AMF40RC60 mixture, which may or may not have had an effect on the ability of PC1 to distinguish between this and the AMF20RC80 mixture. There are no other apparent patterns within this plot.

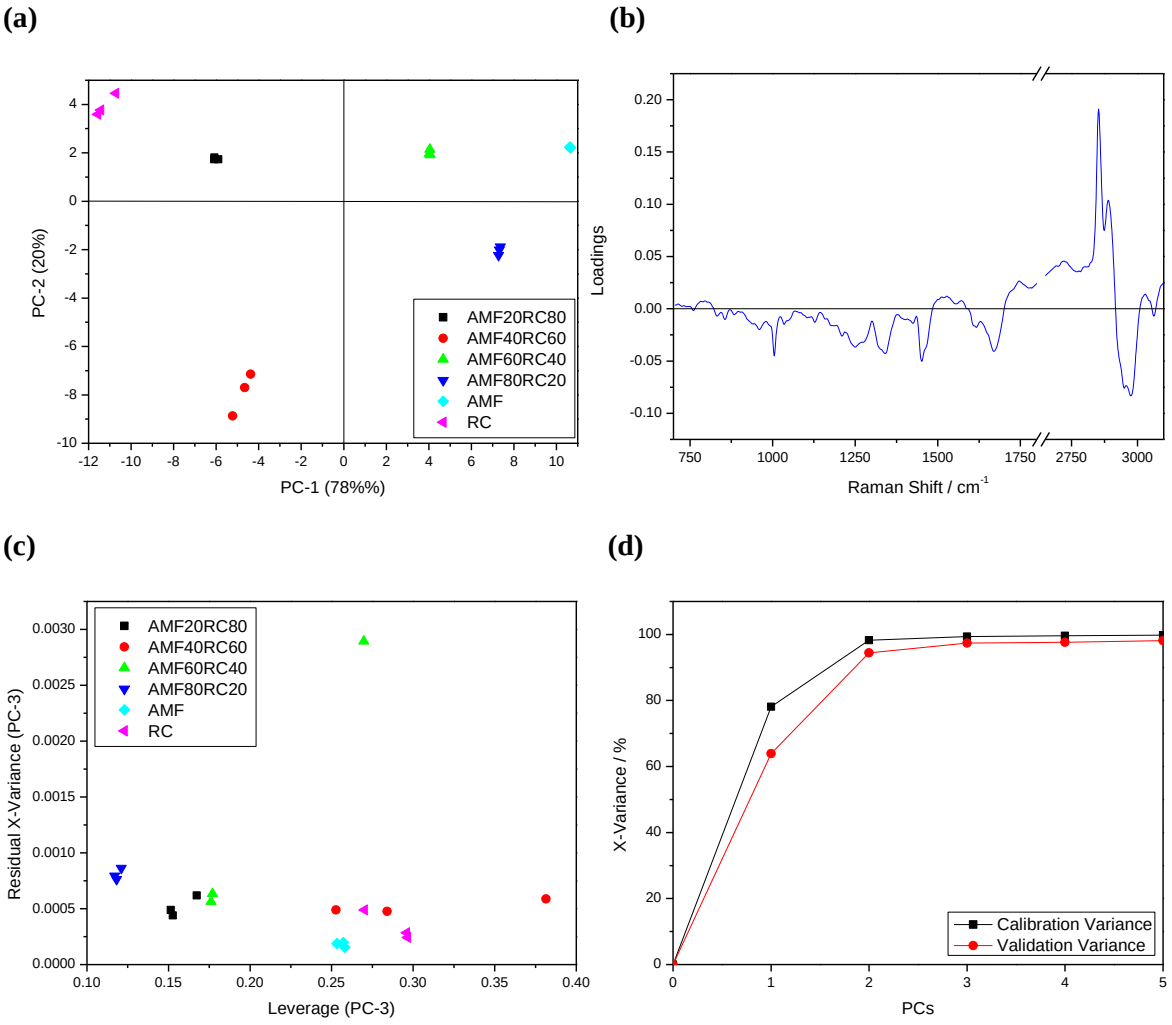



**Figure 5.2 Results of principal component analysis for AMF/rennet casein. This figure contains (a) scores plot, (b) the loadings plot, (c) the influence plot, and (d) the explained variance plot. The keys in the scores and influence plots indicate the symbols which correspond to each sample.**

AMF/SUC mixtures are also very distinct from each other with respect to their grouping on the scores plot (**figure 5.3**). There is also a discernible pattern along the x-axis where the proportion of AMF within each mixture increases from left to right. There is no such pattern with respect to PC2; however, it provides greater distinction between mixtures. This result indicates that AMF/SUC mixtures can be separated based on the component concentrations. The explained variance plot indicates that 97% of the total variance is accounted for using PC1 (85%) and PC2 (12%). These samples, unlike the others, contain several precise regions of variance according to the loadings plot (AMF/CS and AMF/RC both contained very broad

regions). The influence plot does not contain any obvious patterns; however, it does show the possible presence of an outlier from the AMF20SUC80 mixture.

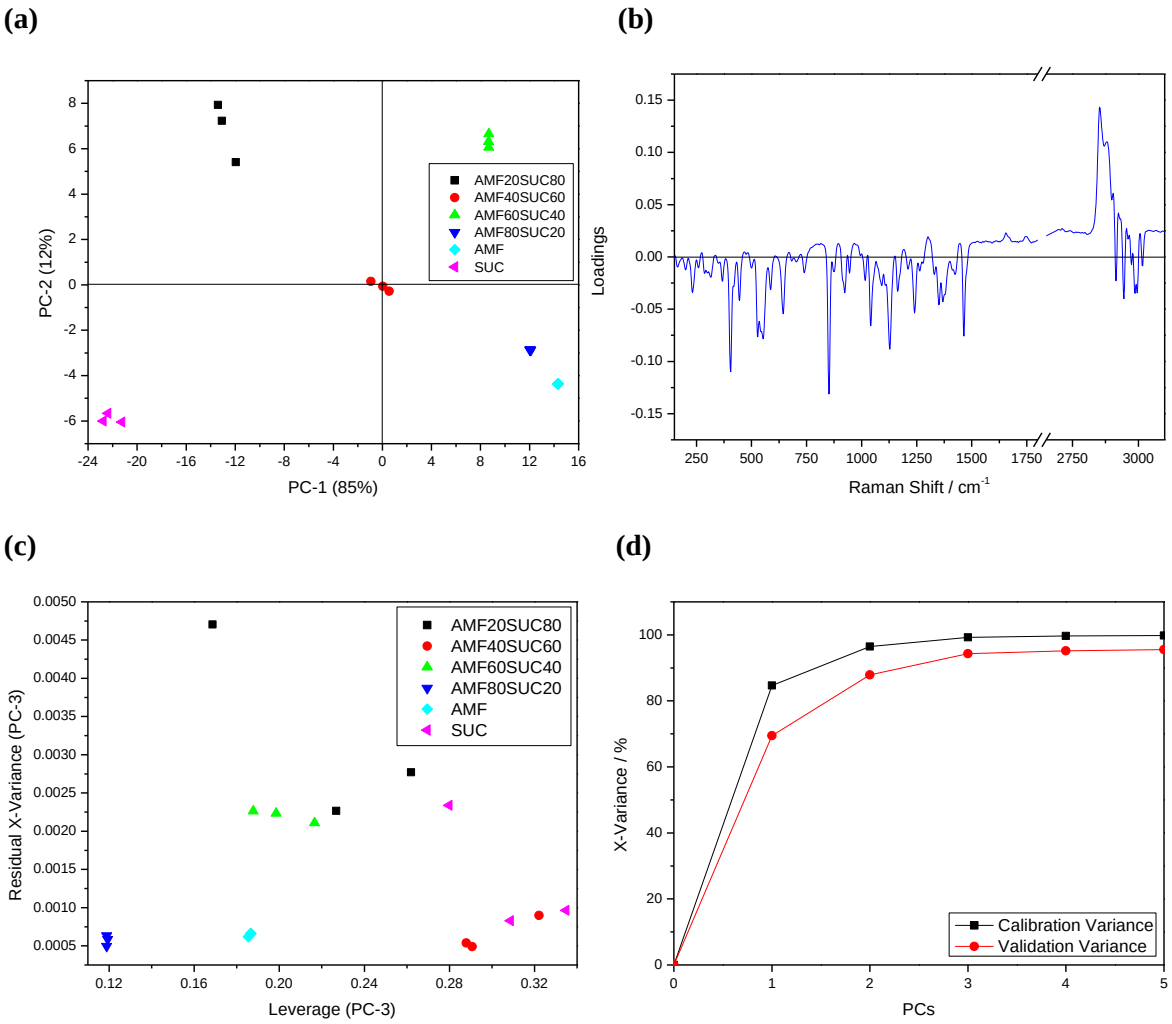



**Figure 5.3 Results of principal component analysis of AMF/sucrose. This figure contains: (a) the scores plot, (b) the loadings plot, (c) the influence plot, and (d) the explained variance plot. The keys in the scores and influence plots indicate the symbols which correspond to each sample.**

## 5.3 Interactions within 1:1 Binary Mixtures

Cheese is a complex mixture of several components including AMF, rennet casein, sucrose, cornstarch and emulsifiers. Therefore, interactions may occur between two or more components. Identification of such interactions was attempted using the Raman microscope with 1:1 AMF/component B mixtures. They would be recognized as band shift or band broadening, in that an interaction would manifest as having bonding character, and therefore

vibrational band structure which was not a simple linear combination of the two component materials. After qualitative analysis of several spectra from each of the binary mixtures, no such spectral regions were identified. The heterogeneity of the samples made the observation of any interface interactions difficult; each sample was generally made up of smooth regions (due to the presence of AMF) and rough regions (due to a granule of component B). For example, in an AMF/SUC mixture, a rough region of the sample is analysed the spectrum will be dominated by the peaks typical of sucrose, whereas a smooth region gives a spectrum dominated by AMF. Therefore the large amount of variation between spectra, solely due to the region analysed, made identification of interactions difficult. The high resolution also hindered identification of interactions; therefore, this was decreased to 9-18 $cm^{-1}$ from 3-5 $cm^{-1}$ to reduce intruding noise in the 2D and 3D maps, discussed in Chapter 6. **Figure 5.4** gives an example of these difficulties.

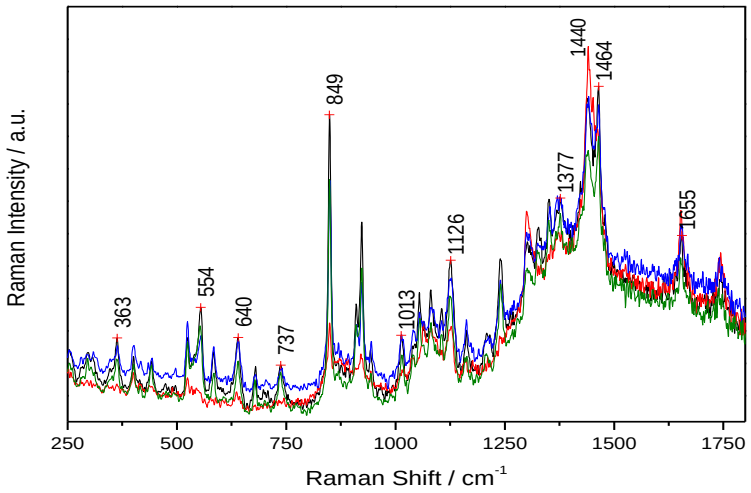



**Figure 5.4 Several spectra taken at various points of the AMF/SUC mixture. This figure shows various peak intensities depending on where the spectrum was taken in the mixture. Raman microscope parameters include: 785 nm wavelength laser, 100 mW power, 50x objective, 25 µm aperture, 6 x 10 s scan time and high resolution.**

## 5.4 Interactions within Tertiary Mixtures

Despite the difficulty of detecting interactions within binary mixtures, heated tertiary mixtures were then analysed involving a 1:1 ratio of AMF with 'component B' and 3% w/w of the emulsifier, trisodium citrate (TSC). The results of this experiment gave spectra which were similar to those collected of the mixtures without TSC. This is thought to be due to the very

small amount of the emulsifier present in the sample and its absence in any of the regions where spectra were taken. However, the use of such a small percentage of TSC in the sample was intentional as this is the proportion used in cheese production. Each spectrum taken from a tertiary mixture showed no direct evidence of the presence of TSC, similar to results shown in **figures 5.5** of a mixture of AMF, cornstarch and TSC. These spectra also show no evidence of interactions between components. The spectra appear to vary only in band height depending on the component content of the region analysed. If component interactions are occurring, they may be more obvious while analysing sample morphology. Therefore in Chapter 6, the morphology of the mixtures was examined via mapping with Raman microscopy for two reasons: (1) to see whether the morphology could be measured and (2) assuming that morphology can be measured using Raman microscopy, could evidence of component interactions then be observed.

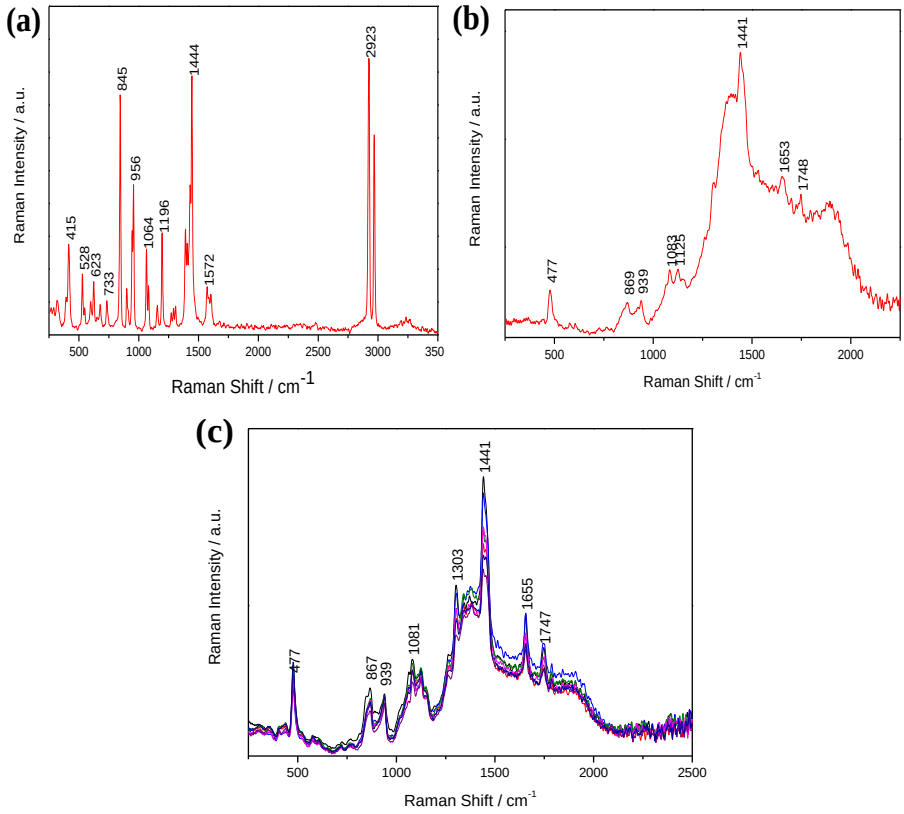



**Figure 5.5 (a) A spectrum of pure trisodium citrate (TSC) taken using the FT-Raman spectrometer. Parameters include: 4 cm$^{-1}$, 264 scans and 120 mW laser power. (b) A low resolution spectrum of AMF60CS40 obtained using the Raman microscope. Parameters include: 785 nm laser wavelength, 100 mW power, 50x objective, 50 µm aperture and 3 x 2 s scan time.**

**(c) Eight spectra of a 1:1 AMF/CS mixture with 3% trisodium citrate (TSC) obtained using the Raman microscope. Parameters include: 785 nm laser wavelength, 100 mW power, 50x objective, 25 µm aperture, 6 x 10 s scan time and low resolution diffraction grating (9-18 $cm^{-1}$).**

## 5.5 Conclusions

This investigation established that bi-component mixtures are able to be distinguished based on relative component concentrations using Raman spectroscopy with PCA. However, identification of interactions between components was not observed because spectra were dominated by features created by the more 'abundant' component within the region. The spectral features were therefore highly variable because of the small measurement spot size which, in turn, hindered identification of interactions. TSC was not able to be found within the tertiary mixtures due to the very low concentration of TSC in each sample and the very small measurement spot size used to collect spectra. It was concluded that if interactions occurred, they were more likely to be observed after mapping the samples.

# Chapter 6 Raman Mapping of Model Cheese Systems

## 6.1 Introduction

Digitised mapping is a useful tool in the imaging of certain samples due to the challenge of discerning the composition of a sample using only optical inspection. Binary mixtures such as those examined in this thesis, were used as a relatively simple model to explore the potential of mapping using spectral data. Two techniques which can be used to study sample morphology are mapping using confocal Raman microscopy and imaging via confocal laser scanning microscopy (CLSM).

CLSM involves staining the samples with dyes which fluoresce when excited by specific laser wavelengths. The different components in a sample can be identified by the dye that stained them. This technique has been used on cheese samples for characterising the distribution of milk fat within a casein matrix.[54] In this investigation the technique was used to view Emmental cheese after four different stages of the manufacture process. A summary of the cheese-making process can be found in **section 1.4**. Pictures of rennet induced curd, heated curd grain before pressing, Emmental cheese after pressing and Emmental cheese after brining were obtained.[54] There were considerable differences in the fat structure between each stage of processing. In the rennet-induced curd the fat globules were small (average diameter: 4.47 ± 0.06 μm) and spherical.[54] Heated curd grains before pressing contained individual, aggregated and coalesced fat globules. The average fat globule size of this sample was ~5 μm.[54] Samples of Emmental cheese after pressing provided evidence of milk fat globule disruption. Cavities containing fat and whey surrounded by protein strands were observed. Fat was found in different forms: milk fat globules similar to those seen in samples after coagulation (with an average diameter of ~4 to 5 μm); aggregates of fat globules and coalesced fat globules; and free fat which is defined as fat not protected by the milk fat globule membrane and susceptible to oiling off or removal.[54] Large fat globules were also found with diameters as large as 40 μm. The average fat globule diameter in this sample was 8.79 ± 0.18 μm.[54] After brining, it was apparent that pools of fat were surrounded by pockets of whey. The protein network was also found to be more continuous and compact than after pressing. Therefore the technique has been found to be useful in the analysis of cheese;

specifically in the analysis of milk fat and rennet casein. CLSM is a fast and effective technique for such systems and has been used extensively by Fonterra in the analysis of cheese products.

Raman microscopy has also been used in an attempt to map cheese to locate the individual components using a combination of univariate and multivariate analyses.[9] In this case, univariate analysis involved integration of a single peak typical of a particular component, using it to identify areas which contain that component and for calculating its relative concentration. This was effective for characterising highly heterogeneous samples such as those with distinct constituent concentration changes within the defined area; however, this method was less useful for discerning component distribution within more homogeneous samples. Multivariate analysis involved the use of PCA and multivariate curve resolution analysis (MCR). In maps created using PCA, the components were spatially separated using scores values. In doing so, several squares were generated within the mapped areas (corresponding to specific components) which were referred to as spatial scores plots. Coupled to each spatial scores plot was a loadings plot. The concentration of a constituent was inferred by the loadings plot.[9] For example, a constituent was regarded as having a high concentration if the spectral regions coincident with the specific constituent were heavily positively loaded. MCR is a technique which may be used to determine the number of constituents within a sample, estimate the spectra of these constituents and to predict the relative concentrations of each constituent.[9, 55, 56] MCR calculated maps were referred to as spatial estimated concentration plots because the values determining pixel colour in the plots were the estimated concentrations. Spatial estimated concentration plots were interpreted by initially identifying the constituent to which the estimated spectrum corresponds. Areas within the map with high component concentration corresponded to localised areas of the constituent identified from the estimated spectrum. Unlike the univariate analyses, the multivariate techniques were found to be reasonably effective for both homogeneous and heterogeneous samples.

However, use of Raman microscopy to analyse cheese products is not common. The benefit of mapping is that no dye is required hence there is no preparation time; no risk of accidentally

altering the sample, such as changing sample consistency; and some components which cannot be stained can give characteristic Raman peaks. One of the aims of this project was to identify whether confocal Raman microscopy could become an alternative for the imaging of cheese and, indeed, other food products also. Model cheese systems containing AMF and component B (cornstarch, rennet casein or sucrose) were investigated in a similar method to McGoverin's univariate method. The aim was to establish a clearer outline of the strengths and weaknesses of the technique for the analysis of cheese samples and to investigate whether three-dimensional contour maps can be generated and how these maps compare to similar images obtained using CLSM.

## 6.2 Results and discussion of 2D contour maps

The contour maps illustrate the relative intensities of specific peaks within a pre-defined area of a sample. The peak intensities, to which this thesis refers, are actually the integral values obtained by calculating the area underneath the respective peak. Larger constituent quantities generally produce larger peak heights (peak intensities) which results in larger integral values obtained from underneath peaks. The largest peak intensities are represented as the red regions in the contour map while the least intense peaks are represented by the blue/black regions of the contour map. The different peak intensity 'levels' (represented by the different colours) are assigned to be relative to the *same* peak in the *same* sample. For example, the red region for RC in AMF60RC40 does not necessarily indicate the presence of the same peak intensity as the red region in AMF80RC20, nor does the same colour in both AMF60SUC40 and AMF60CS40 indicate the same intensity of AMF peak. The implications of the lack of translatability between sample contour maps are discussed in **section 6.4**. The relative intensity levels are illustrated in **figure 6.1** and an example of raw contour maps (of AMF80RC20) and a combined component contour map made using Adobe Photoshop CS 5.5 (Adobe Systems Inc., San Jose, U.S.A.) in **figure 6.2**. Each contour map contains the **same scale** for each of the axes (i.e. spanning 130 μm for the x axis and 100 μm for the y axis). Units for each axis are in micrometers (μm).

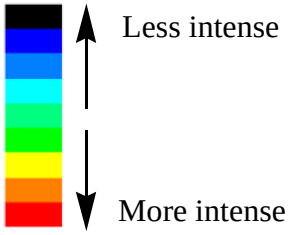



**Figure 6.1 A diagram illustrating the ranking of intensity levels.**

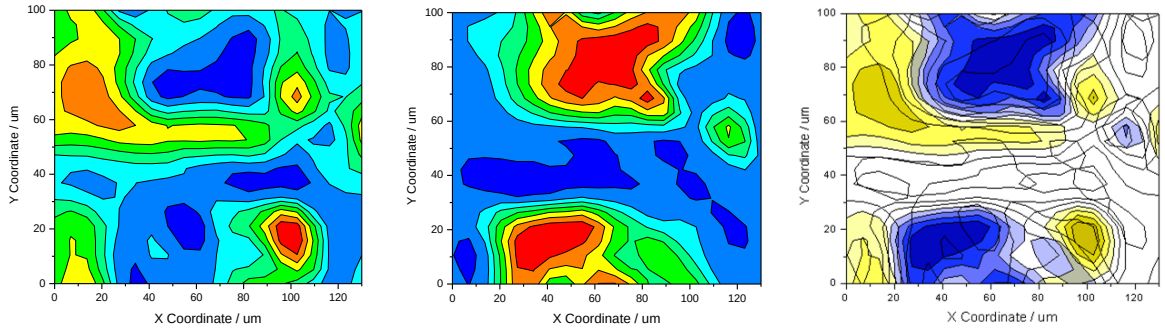



**Figure 6.2 Three contour maps from AMF80RC20. This includes a contour map of AMF (left), RC (middle) and a combined component contour map (right). The dark regions in the contour map on the right indicate regions containing more intense peaks. Intensity thresholds were kept the same as those that Origin 7.5 initially calculated; the contour map on the right only contains the 3 most intense thresholds for each component and was created using Adobe Photoshop CS 5.5 (Adobe Systems Inc., San Jose, USA). The units for the X and Y coordinates are in micrometers (μm).**

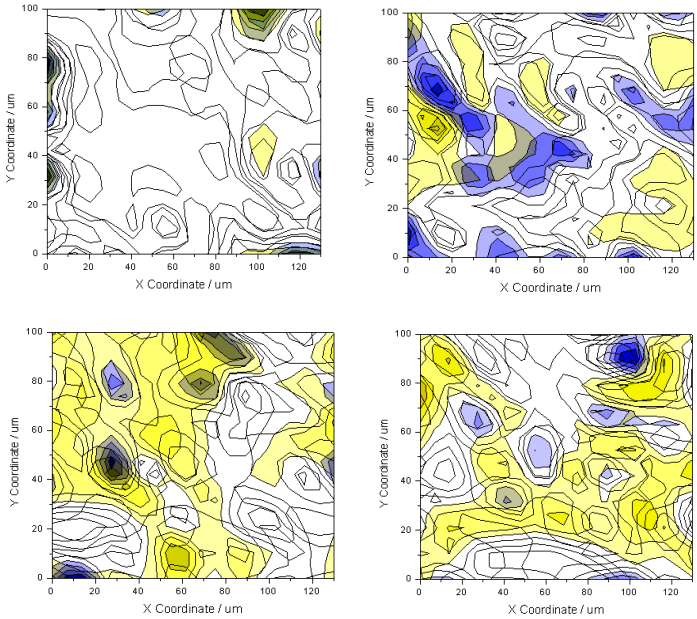



**Figure 6.3 Contour maps of samples containing AMF and cornstarch. AMF20CS80 (top left), AMF40CS60 (top right), AMF60CS40 (bottom left), and AMF80CS20 (bottom right). Darker regions indicate regions of higher peak intensity. Intensity thresholds were the same as those that Origin 7.5 initially calculated; only the 3 most intense thresholds for each component were used.**

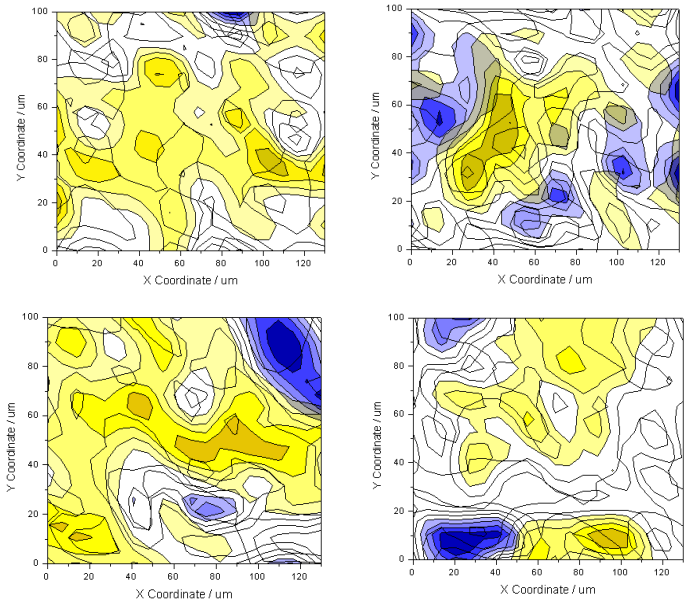



**Figure 6.4 Contour maps of samples containing AMF and sucrose. AMF20SUC80 (top left), AMF40SUC60 (top right), AMF60SUC40 (bottom left), and AMF80SUC20 (bottom right). Darker regions indicate regions of higher peak intensity. Intensity thresholds were the same as those that Origin 7.5 initially calculated; only the 3 most intense thresholds for each component were used.**

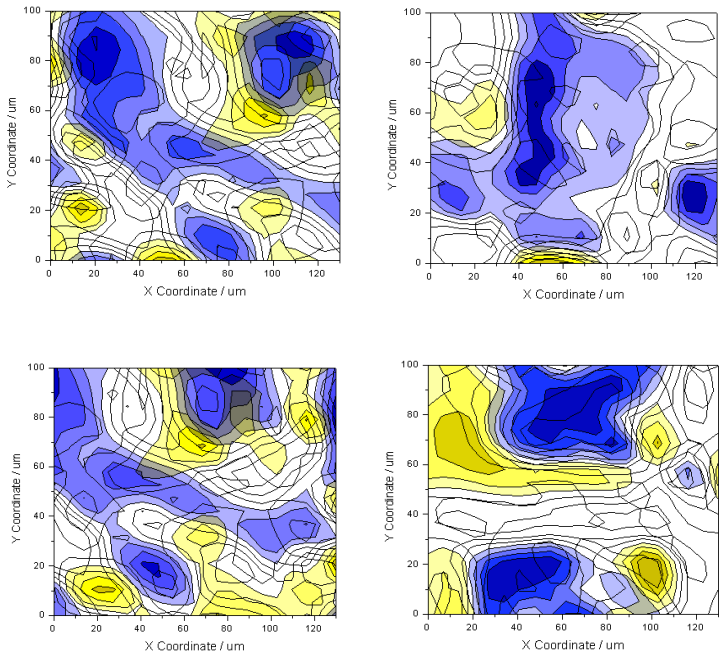



**Figure 6.5 Contour maps of samples containing AMF and rennet casein. AMF20RC80 (top left), AMF40RC60 (top right), AMF60RC40 (bottom left), and AMF80RC20 (bottom right). Darker regions indicate regions of higher peak intensity. Intensity thresholds were kept the same as those that Origin 7.5 initially calculated; only the 3 most intense thresholds for each component were used.**

**Figures 6.3, 6.4** and **6.5** illustrate that the contour maps from both AMF and the second component tend to be able to be combined with minimal overlap occurring between the two components. Exceptions include two of the AMF/CS contour maps and two of the AMF/RC contour maps. AMF/RC and AMF/SUC contour maps show the strongest tendency for mutual exclusion between the individual components. This means that if there is a region of the map indicating the presence of one component, the other component is shown to be absent in the region, but rather located in close proximity surrounding the other component instead. AMF/CS does not seem to do this in all cases. The overlap between the components could be indicative of interactions between the components or simply be the result of the poor resolution of the microscope in the z-axis and the signal of a granule becoming mixed with AMF above or below it. AMF20CS80 in **figure 6.3** can be seen to involve a large amount of overlap relative to the very small amount of signal present. Poor microscope focus is believed to be the cause of the poor signal. Taking these results into account it was believed the 3D

mapping will be possible and would provide more insight into the ability of Raman microscopy to image cheese.

Fluorescence was a factor which affected the spectra used to create contour maps and was caused by the glass slides upon which the samples were placed. **Figure 6.6** illustrates the region in which fluorescence is visible. This had a tendency to occur in regions where the sample was smeared thinly on the slide. The effect of this was buffered by changing the integration method from integrating the entire area under the peak to integrating the area within the peak by essentially connecting the two points either side of the peak by a straight line and finding the area below the peak but above the straight line. It was found that attempting to baseline correct the spectra caused the peak heights to artificially be adjusted and not necessarily according to the height of the peak caused by fluorescence. Using the different integration method was found to be a reliable method to negate the effect of the fluorescence without significantly affecting the results in an adverse manner. This is discussed later in this chapter.

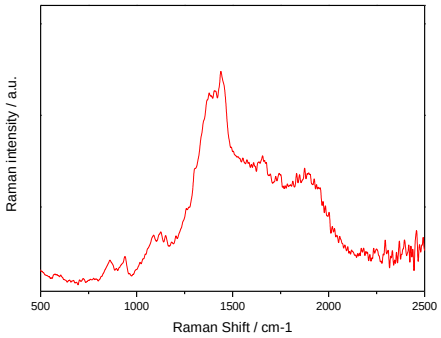


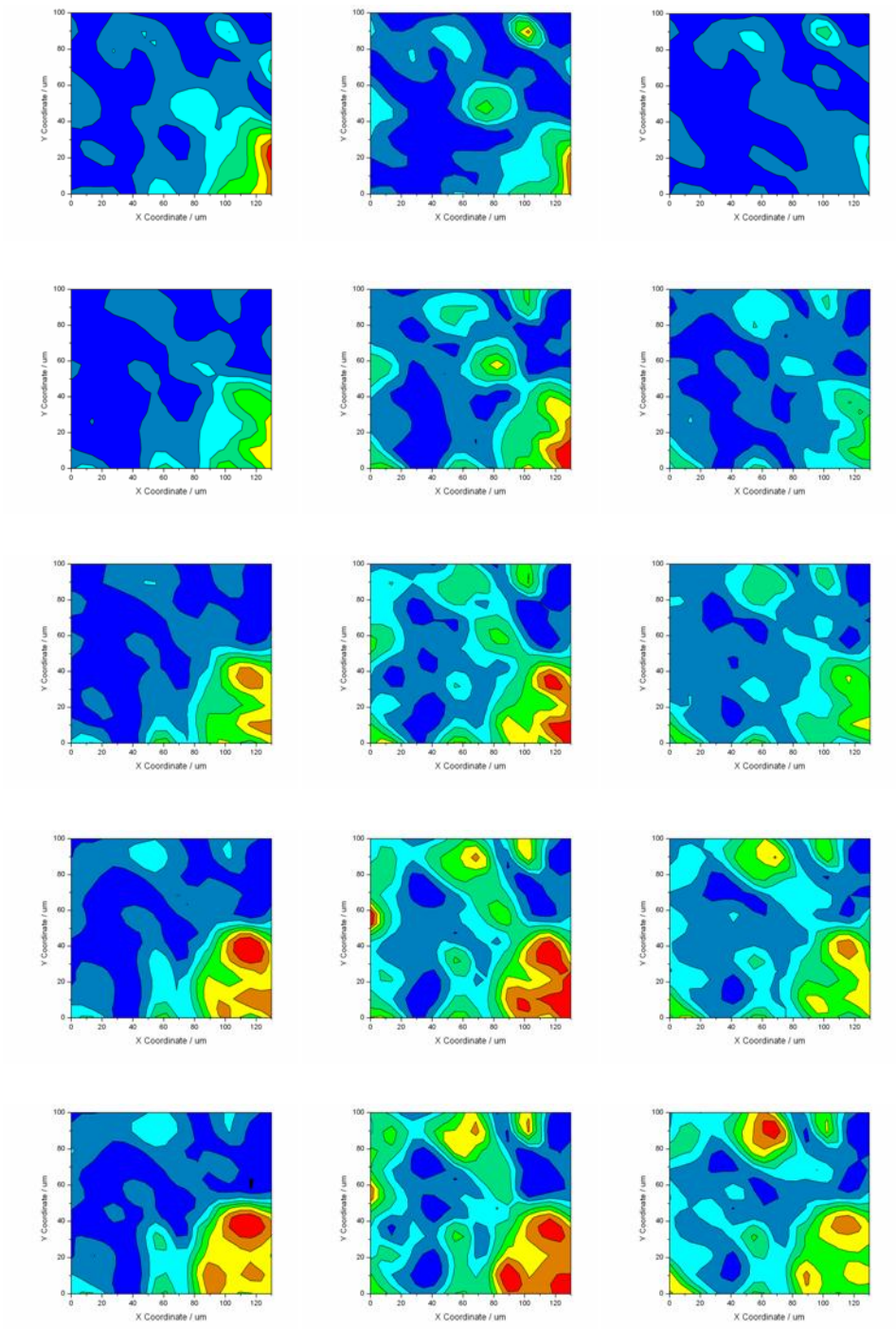

**Figure 6.6 Spectrum of the fluorescence from glass slides between 1250 $cm^{-1}$ to 2000 $cm^{-1}$.**

## 6.3 Results and discussion of 3D contour maps

From the results of the 2D mapping project, it was be seen that contour maps could indeed be constructed from the integration of individual peaks. However, reliability of the technique was not certain. Therefore, with the purpose of testing both the reliability of the contour maps and the ability of the Raman microscope to perform the same imaging tasks as CLSM, 3D mapping was attempted. However, the results of the 3D mapping will be discussed primarily with respect to the sequences of 2D contour maps from which they are made. This is because the 2D contour map sequences are much easier to refer to and they illustrate the component

distribution in the centre of the 3D maps much more clearly. These maps will be discussed with respect to how consistent each map is in relation to the maps that surround it. Also, the 3D contour maps obtained for each mixture will be compared to those obtained during the 2D mapping to see whether there are any inconsistencies between results. Furthermore, maps were created using different peaks from the same component in order to determine how reliably the peaks indicate the presence of the respective component. The contour maps present in **figure 6.7** are the result of the 3D mapping. They are separated into columns according to which peak in the spectrum they were created from. The rows in each column represent the different layers within the 3D contour map. The top contour map in the column is the bottom layer of the sample and each layer is 3.2 μm higher than the previous layer (see Appendix A).

The 3D mapping results suggest that the contour maps can indeed reliably indicate the location of each component within the samples. Ensuring this was the case was achieved through analysis of two factors: the change between layers and the use of different peaks to create the maps. Contour maps created using peaks from sucrose in the AMF60SUC40 mixture are presented in **figure 6.7**. For clarity, notes about which component and which sample are relevant to each figure will often accompany a figure number in this chapter. An example is: **figure 6.7 (sucrose peaks AMF60SUC40)** which indicates that this figure illustrates the contour maps created from sucrose peaks collected from the AMF60SUC40 mixture. Due to the large amount of space used by such figures, the remaining contour map sets for the other components and mixtures are presented in **Appendix B**. Each of the sets of data in **figures 6.7** and **figures** 8**.1** to 8**.5** from **Appendix B** show that each adjacent layer within a sample gives a very similar contour map. It is only until a holistic view of the sample data is obtained before the pattern of gradual changes in the data becomes apparent. Gradual changes between maps suggest reliability in the data because each map consistently indicates the components are in the same location. It is also what one would expect within a sample. For example, if one was to slice a sucrose granule into 10 pieces one would expect the diameter of the top and bottom slices to be the smallest and the diameter to be largest in the centre. This is precisely what was observed when looking at the AMF/SUC sample and each other sample also.


Y Coordinate / um
X Coordinate / um
0
20
40
60
80
100
120

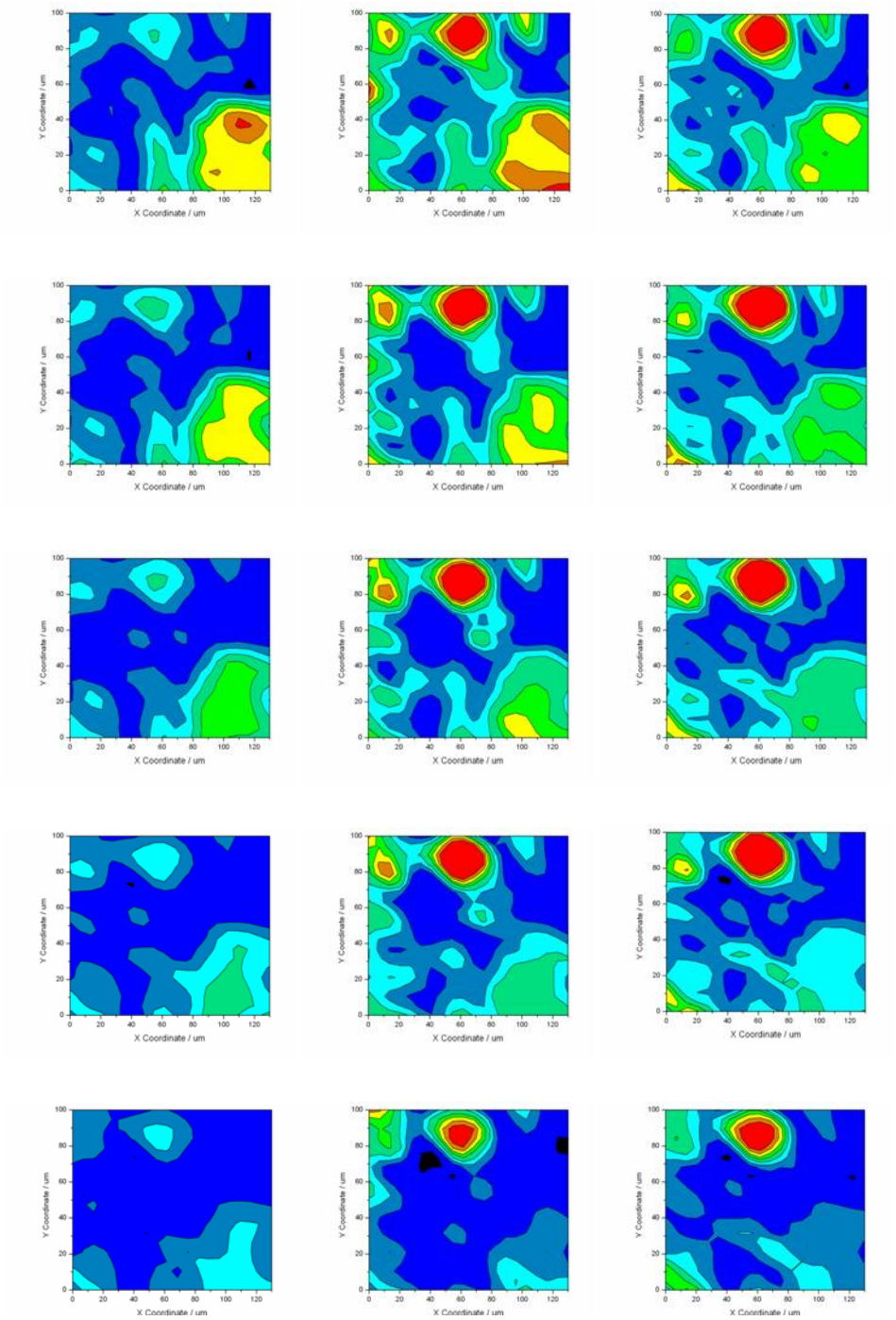



**Figure 6.7 Three columns of contour maps for AMF60SUC40 mixture. Each column of 10 contour maps has been created using a different peak with the top contour map of each column being the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer. All three peaks used are characteristic of the presence of sucrose. The columns consist of: contour maps created using the 539 $cm^{-1}$ peak (left), maps created using the 641 $cm^{-1}$ peak (middle), and maps created using the 849 $cm^{-1}$ peak (right).**

Integrating different peaks to create the contour maps also emphasises the reliability within the results. Each component (with the exception of rennet casein) produces several peaks at Raman shifts which the other components do not. For the 2D contour maps only one of the characteristic peaks of each component was used to indicate its presence in the sample. In order to test how reliable the chosen peaks were, several more sets of contour maps were created using other such peaks. **Figure 8.2 (cornstarch peaks in AMF60CS40)** from **Appendix B** shows how the three peaks used to indicate the presence of cornstarch all give very similar contour maps for each corresponding layer within the sample. This suggests the peaks chosen are likely to reliably indicate the presence of cornstarch. All samples gave this same result for each peak chosen except the 1655 $cm^{-1}$ peak in AMF60RC40 which was found to give maps significantly different to those of all other peaks chosen for AMF. The 1300 $cm^{-1}$ peak was used in this case in order to establish which contour maps were more likely to be illustrating the locations of AMF's presence. It became apparent that the 1655 $cm^{-1}$ peak was the most likely to be producing the erroneous result. The likely reason for this was a pair of broad peaks at 1668 and a 1616 $cm^{-1}$ caused by RC. Therefore the contour maps created using the 1655 $cm^{-1}$ peak were likely to be inadvertently indicating the presence of AMF in situations where there was actually an absence of the fat. The reason why RC had only the 1003 $cm^{-1}$ used to indicate its presence was simply because the only other peaks which were considerable was that at 1668 $cm^{-1}$ and one at 1450 $cm^{-1}$ which were both broad, the definition of the peaks was sometimes difficult, and AMF produces stronger peaks in similar locations. The 1440 $cm^{-1}$ peak used as a characteristic peak for AMF can be seen to be affected slightly by the presence of the 1450 $cm^{-1}$ peak as the regions shown to contain no AMF by peaks 1744 $cm^{-1}$ and 1300 $cm^{-1}$ appear to contain small amounts of fat. However, in this case there is a distinct difference (intensity threshold colour) between the regions which 'actually' contain AMF and the regions which 'do not'. This is one of the more obvious flaws in this method of imaging. How can the technique be applied in actual cheese samples if the maps become unreliable as soon as there is overlap of any peaks? Certainly, this technique would be best suited to samples which contain components which produce characteristic peaks which do not overlap. However, applying the technique to cheese samples will become much more difficult as overlap is common between components. Based on spectra of the four components used in this project, identification of AMF, cornstarch and sucrose will be relatively simple as each

produces at least one peak at a unique Raman shift; 477 $cm^{-1}$ for cornstarch, 1744 and 1300 $cm^{-1}$ for AMF and 401, 526, 641and 1239 $cm^{-1}$ for sucrose. RC on the other hand, will be much more difficult to identify, as the only peak which is likely to be able to be used (1003 $cm^{-1}$) is slightly overlapped by a peak at 1013 $cm^{-1}$ from sucrose and overlapped by the broad shoulder of a set of peaks from cornstarch. The exact significance of this is not yet known but can be tested in the future using model systems containing all of these components. In order to devise a possible method to remedy this problem, it was decided to test the effect of adding the integration values of the characteristic peaks for each component. The entire set of results of this test can be found in Appendix B. In **figure 8.6 (combined AMF peaks in AMF60RC40)** from **Appendix B,** the sum of the 1440, 1655 and 1744 $cm^{-1}$ peaks can be seen to have been used to create a new set of contour maps. This result is particularly important as it involves the use of the 1655 $cm^{-1}$ peak which was found to give an erroneous result due to overlap with broad peaks from RC. The combination of the peaks produced a set of contour maps similar to that of the 1440 $cm^{-1}$ data. Despite the use of the erroneous data, the regions containing AMF and those without are once again distinct. Therefore it is possible that this method could be used to reduce the effect of overlap between peaks. This could also be tested in model systems containing all four of the components. As expected, results of the other sets (**figures 8.7** and **8.8**) of combined peak data can be seen to produce contour maps very similar to those created using the single peaks. Once the reliability of the results had been established, each mixture was analysed with respect to the way in which the components mixed and/or interacted.

**AMF60SUC40**

Contour maps for each mixture (AMF60SUC40, for example) can be split into two categories: those which were made using peaks of AMF and those which were made using peaks of SUC. Each column in each figure is a separate set of contour maps which were created using different peaks of the **same component**. For example, **figure 6.7** contains three columns; the left column was created using the 539 $cm^{-1}$ peak, the 641 $cm^{-1}$ peak was used to create the middle set of contour maps and the 849 $cm^{-1}$ peak was used to create the right column. Each of these peaks is cause by the presence of sucrose. From the each figure it can be seen that the 539, 641 and 849 $cm^{-1}$ each gave very similar-looking sets of contour maps. The 1440, 1655

and 1744 $cm^{-1}$ peaks did also. This suggests that each peak is indeed indicating the true location of the respective component. However, this sample result did give a subtle indication that it too was affected by the overlap of characteristic peaks. This is evident at the top of each of the contour maps in **figure 6.7 (sucrose peaks in AMF60SUC40)** and **figure 8.1 (AMF peaks in AMF60SUC40)** from **Appendix B**. All sucrose contour maps show the presence of a relatively intense circular region near the top of the map while all but the 1440 $cm^{-1}$ AMF column of contour map show a relatively weak (circular shaped) signal in this region. Instead, the 1440 $cm^{-1}$ peak gives a relatively intense region in this location. This is likely to be due to the presence of peaks at 1463 and 1429 $cm^{-1}$ from sucrose. Interestingly, these peaks affected the top middle region of the map yet the bottom right hand corner of the map (which is the location of another very intense sucrose signal) remained with relatively weak signal around the 1440 $cm^{-1}$ region. The reason for this is still unknown. Other than the aforementioned anomaly, AMF60SUC40 gave contour maps which generally contained component signals which were mutually exclusive with respect to their map location. This means that where there was a region of intense signal from sucrose there would be a relatively weak signal from AMF. This is in agreement with the 2D mapping result for AMF/SUC mixtures which further supports the reliability of the methodology. The AMF60SUC40 3D map, which can be found as a video clip on the supplementary DVD, was constructed using the 1655 $cm^{-1}$ and 641 $cm^{-1}$ data sets. The resulting video shows the 3D spatial arrangement of both components which and clearly shows the mutually exclusive nature of the mixture. The sample can be seen to look like granules suspended in a viscous liquid medium.

**AMF60CS40**

The series of 2D contour maps created using this sample are present in **figures 8.2 (CS peaks in AMF60CS40)** and **8.3 (AMF peaks in AMF60CS40)** from **Appendix B**. The 3D images and the series of the combined component 2D contour maps can be found on the DVD. This sample had 12 layers of contour maps created because of the lack of signal found for the bottom two layers. Unlike the other samples, AMF60CS40 gave contour maps where the locations of the components were not mutually exclusive with respect to each other. The presence of the components appears to be independent from whether or not the other component is present. This is most evident from the 2D series on the DVD. Interaction

between the components could be a possible reason for the abundance of overlap. The most likely way in which this would take place is if the AMF was dissolving the cornstarch. This is not unexpected as the non-polar nature of both components would facilitate this. However, the spectra obtained from the regions of overlap gave no supporting evidence (such as band broadening or shifting). Also unlike the other samples, the overlap in the contour map cannot be explained by peaks sharing similar Raman shifts as both the 1655 $cm^{-1}$ and 1744 $cm^{-1}$ peaks are both located a large distance from any cornstarch peak –likewise for all CS peaks relative to AMF peaks.

**AMF60RC40**

**Figures 8.4 (AMF peaks in AMF60RC40)** and **8.5 (RC peaks in AMF60RC40)** from **Appendix B** both illustrate the separate component series of contour maps. Note that for this mixture, only one set of contour maps (using only the 1003 $cm^{-1}$ peak) was created for RC. This is due to RC peaks being relatively broad and weak compared with AMF peaks and peak overlap between components. The 1003 $cm^{-1}$ peak was the most well defined peak of RC. The 3D image and combined series of contour maps on the DVD allow the extent of mutual exclusion to be observed more effectively. This sample is notable for the very well defined component locations. Observed separately, one can see the RC contour maps clearly show at least two RC granules on the left hand side separated by a small gap, while on the right hand side, part of several granules can be seen. The centres of the maps indicate a consistent absence of RC. Combining the maps for both components clearly explains the absence as the centre is almost completely filled with AMF. The very precise location of the sliver of AMF between the two large RC granules attests to the accuracy of the technique for mapping. This sample is also notable for the apparent lack of signal for both components in the centre of the maps; a result which appears to be replicated by all the peaks analysed. **Figure 6.8** is a picture taken using the microscope camera and distinctly shows the presence of a similarly sized object in the same location. However, the lack of any Raman signal in this area hindered attempts to identify the object. It is likely that the lack of signal is due to an air bubble, which was commonly found in samples containing larger concentrations of AMF.

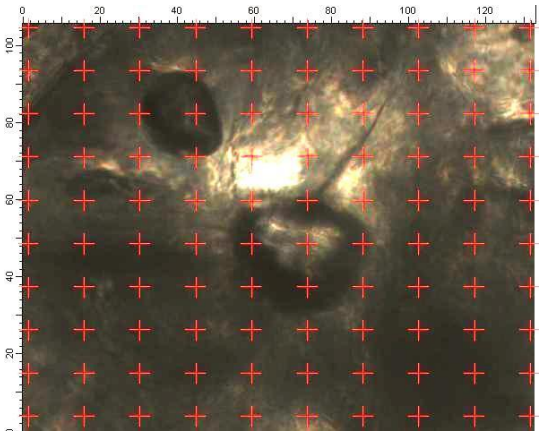



**Figure 6.8 Picture of AMF60RC40 layer 5, taken using the Raman microscope camera (units: µm).**

## 6.4 Alternative Standardisation Method

In addition to using the grey scale method of combining contour maps, another method was used involving utilisation of a threshold peak intensity to determine the precise borderlines of each component. The reason for employing such a method is to observe whether or not the technique has comparable x/y resolution with confocal laser scanning microscopy and to explore whether a standardised method could be applied to all samples rather than using a technique where the maps created use intensities which are based solely on the relative intensities within the sample. Use of thresholds to decide the presence (or absence) of a substance can be found in literature for use in the mapping of pharmaceutical drugs in solid lipid matrices.[8] At each point on the contour map, several peaks are integrated and from this it is determined whether or not that point primarily contains AMF or component B. The goal of applying a threshold value for each peak is to be able to say that above a certain threshold value the location on the contour map contains the component and integration values below the threshold indicate the absence of the component. For example, looking at the 1440 $cm^{-1}$ peak for AMF on a reference spectrum, we find integrating this peak gives a value of 'X'. We then decide that the threshold value for the presence of AMF in all samples is '10% of X'. Upon integrating all other samples at peak 1440 $cm^{-1}$ we decide that if the value calculated is above '10% of X' then AMF is present and if it is below '10% of X' then there is no AMF present. Appropriate thresholds were sought via qualitative evaluation of the effects on the contour maps when the threshold was changed. Thresholds of 10%, 12.5%, 15% and 20%

were tested for AMF60RC40 contour maps (since contour maps from this sample fit well). The 12.5% threshold produced maps where the components fitted with minimal overlap and/or gaps. This process is illustrated in **figure 6.9**.

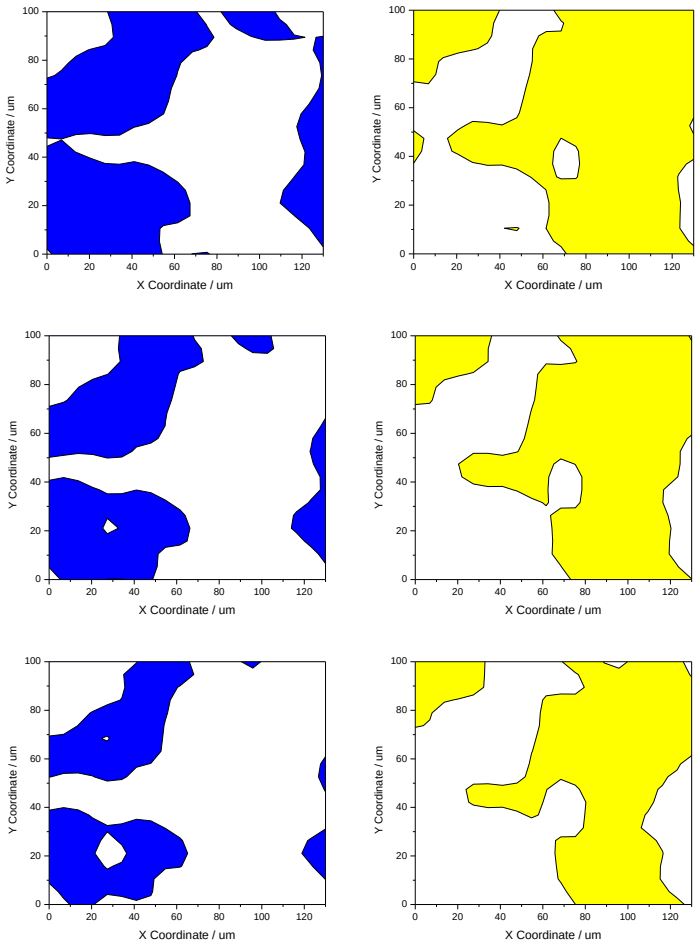



**Figure 6.9 Column of RC contour maps AMF60RC40 layer 8 (left), column of AMF contour maps AMF60RC40 layer 8 (right). Top row contour maps were made using a 10% threshold value. Middle row contour maps were made using 12.5% threshold value. Bottom row contour maps were made using 15% threshold value.**

From **figure 6.10** it can be seen that even though a threshold value was found which could be applied to AMF60RC40, applying the same threshold value to different samples gives erratic contour maps. Using the same threshold for the sample that contains sucrose results in large amounts of overlap due to signal from sucrose essentially flooding the map. Conversely, signals from both AMF and CS in AMF60CS40 are scarce and the combined map remains

largely empty. Using the same threshold for different components appears to be unsuitable because the introduction of other substances affects each peak differently. Indeed, even the same peaks had differing relative intensities when mixed with different components. Therefore it is evident that this particular process of standardising the methodology cannot be used.

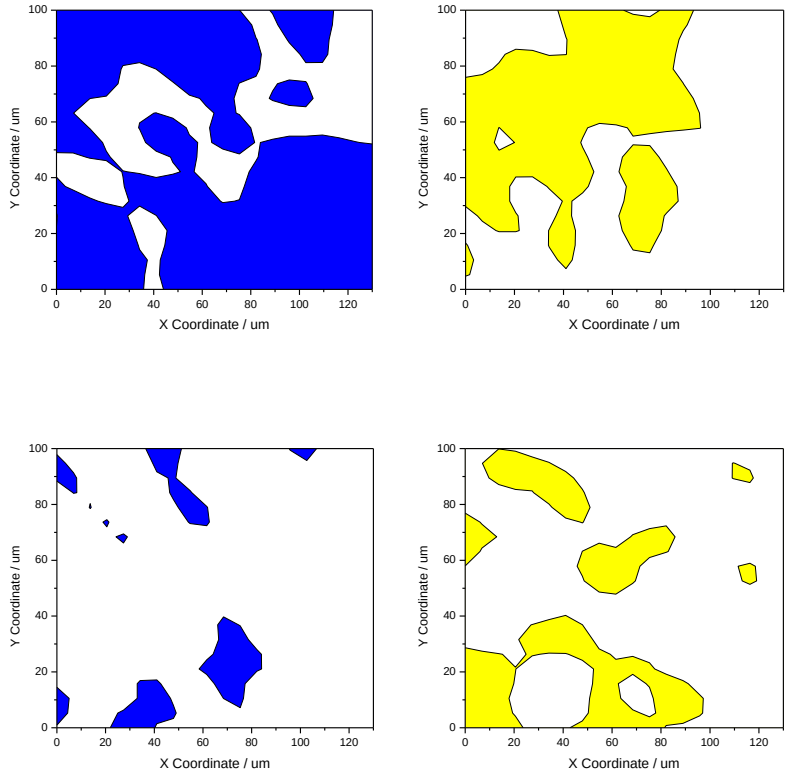



**Figure 6.10 AMF60SUC40 contour maps of the 539 $cm^{-1}$ and 1440 $cm^{-1}$ peaks, respectively (top row), and AMF60CS40 using 940 $cm^{-1}$ and 1655 $cm^{-1}$ peaks, respectively (bottom row). Both sets of contour maps were created using the 12.5% threshold value calculated from each of the respective reference spectra (not shown).**

## 6.5 High Resolution Contour Map

In an attempt to assess the effect various parts of the methodology had on the results, a set of high resolution (HR) contour maps were created. AMF60SUC40 was used because it appeared to have the highest chance of containing overlap but also demonstrated a certain degree of mutual exclusivity for the location in which each component resides. The differences in the

methodology included changing the number of points on the 130 µm x 100 µm grid from 10 x 10 to 30 x 30. In order to avoid the extremely large acquisition time which would have resulted from taking 900 spectra, the scan time was reduced from 2 x 3 s to 2 x 1 s. Finally, the slide and coverslip used were both made from quartz in order to avoid any unwanted fluorescence. 900 spectra were collected in approximately 4 hours and in a process identical to that used for constructing previous maps, two high resolution contour maps were created (one for each component). The combined component map (**figure 6.11**) was made by Liz Nickless using Image-Pro Plus (version 7, Media Cybernetics, MD, USA). This map, like many of the AMF/SUC maps, illustrates a tendency for mutual exclusion for each of the components with a small amount of overlap. This would indicate that much of the overlap is likely to be caused by the z-resolution of the instrument rather than poor x/y resolution caused by the number of measurement points in the grid. The main difference between AMF/SUC contour maps is simply the apparent texture of the AMF. 100 point maps indicated AMF had a smooth paste-like texture, whereas the 900 point map seemed to show it having a more mottled texture. Despite still containing a large amount of overlap, this result indicates that high resolution maps can be used for identification of component locations and textures; and can be obtained in relatively short time frames.

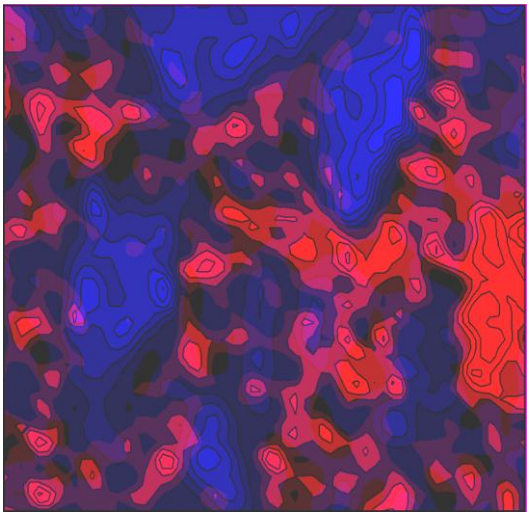


**Figure 6.11 High resolution contour map of AMF (red) and sucrose (blue). This map was made using a 30 x 30 point grid, rather than the 10 x 10 point grid used in previous maps.**

## 6.6 Confocal Laser Scanning Microscopy (CLSM) Results and Discussion

The z-series of AMF60RC40 which was obtained in Dunedin at The University of Otago Anatomy department can be found in the 'CLSM' folder on the supplemental DVD. Three z-series' were created using different regions of the sample. Red regions indicate the presence of AMF due to fluorescence from Nile red while green regions are caused by fluorescence from fast green and are intended to indicate the presence of RC. However, each of the z-series obtained contain green regions which appear to indicate the RC is in liquid phase. These regions were often small and spherical or large and apparently porous; neither attribute is typical of rennet casein. Consistency of the sample as it was placed on the slide would suggest the presence of granular objects immersed in a fatty liquid, which was reinforced when looking at the sample through the microscope objective. Care was taken to create images of the granular regions, yet in each z-series the green regions remained liquid-like. Instead large cavities can be seen in the location where granules were observed. This suggests the failure of the fast green dye to stain the rennet casein. Comparison of the results with images in **figure 6.12** obtained by Elizabeth Nickless support this conclusion. In **figure 6.12** rennet casein appears to take on a granular appearance, shaped similarly to the cavities present in the obtained results and unlike the green regions. AMF can be seen as the red regions in both the results and the reference image in **figure 6.13**. As expected, the AMF appears to resemble a liquid-like substance in both images.

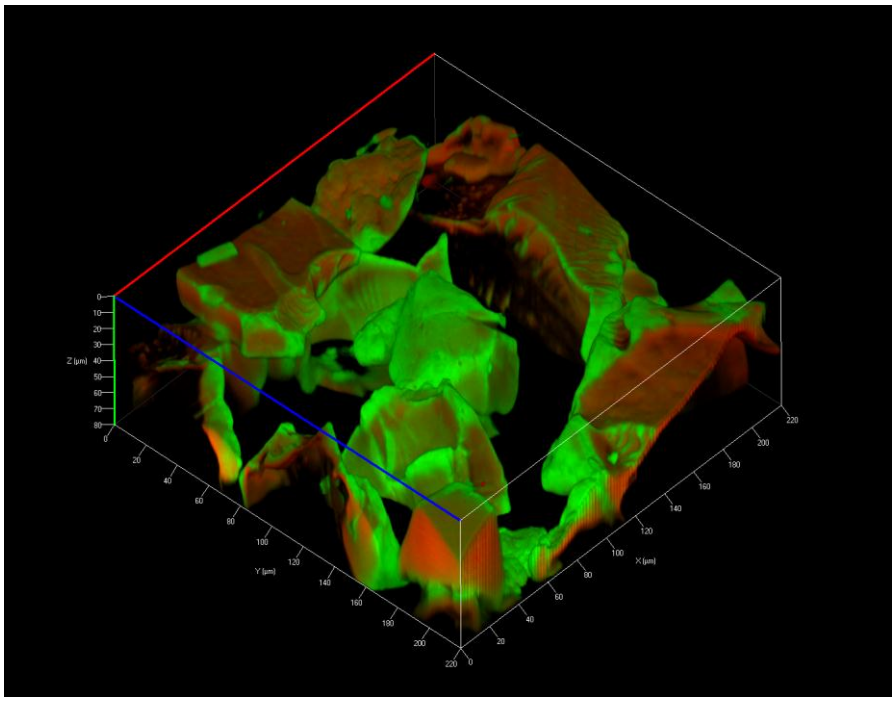


**Figure 6.12 Image of Rennet casein using confocal laser scanning microscopy, provided by Elizabeth Nickless.**

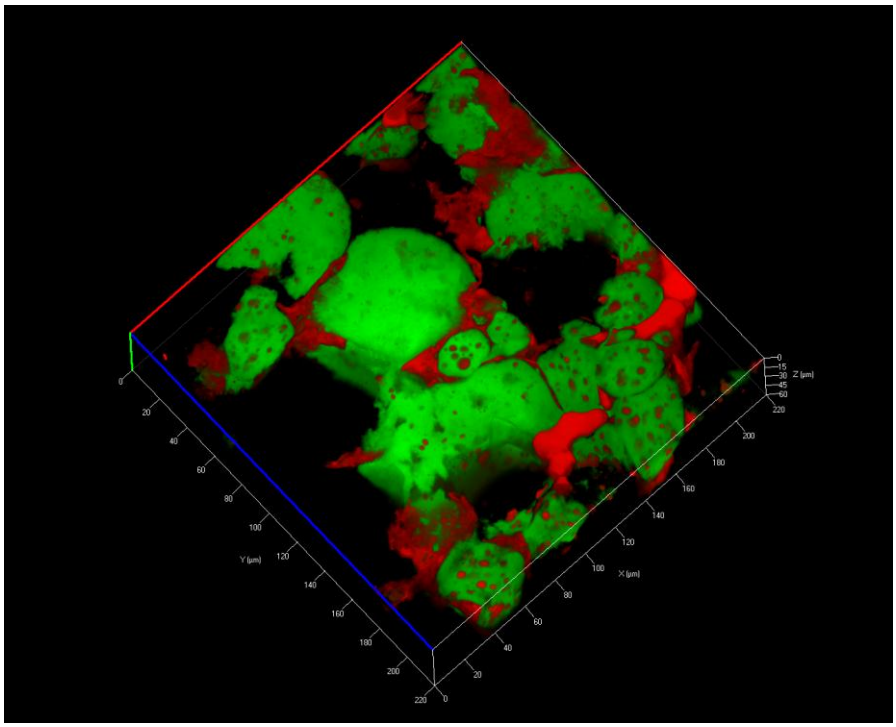


**Figure 6.13 Image of AMF (red) mixed with protein (green), provided by Elizabeth Nickless.**

## 6.7 CLSM versus Raman Microscopy

These two techniques produce images where the relative sizes of cornstarch and rennet casein granules are quite similar. Such consistency attests to the ability of Raman microscopy to accurately image the contents of a sample. There are discrepancies between the techniques when imaging sucrose, however. Possible causes for this include the difficulty of imaging carbohydrates using CLSM and/or the preparation of the binary mixtures. An example of difficulty in staining a sample can be seen from the AMF60RC40 sample discussed in **section 6.6**. Fast green was unable to stain rennet casein and large cavities were observed which were surrounded by large amounts of AMF. However, distinct green lines can be seen around the edges and in the centre of the cavities and suggests that the corners of the granules were able to be stained successfully. This is seen again in the RC image taken using CLSM where the corners of each granule were clearly visible due to the presence of large amounts of stain in these regions. Sucrose could not be stained and the image seen in **figure 6.14** is simply a reflection image obtained by measuring the amount of laser light reflected off the surface of the crystal. It is relatively difficult to distinguish individual crystals, although subtle lines which are likely to be the sharper corners of the crystals can be seen. The lines also orientate to create possible outlines of crystal structures. Broader regions of signal are likely to be rough surfaces of the crystal parallel to the plane of the page. After comparing Raman microscopy and CLSM results, a discrepancy in the size of sucrose granules was noticed. The preparation of samples analysed using the Raman microscope may have been a possible cause. The preparation of the binary mixtures involved geometric dilutions of the components using a mortar and pestle. While RC and CS were able to be mixed with the granules apparently remaining the same size, the large sucrose crystals were significantly altered by the mixing. This was more apparent after preparing the pure samples which were all ‘mixed’ using the mortar and pestle in order to keep the treatment of the samples as similar as possible. After 30 s of mixing, sucrose crystals invariably took on more of a powder consistency.

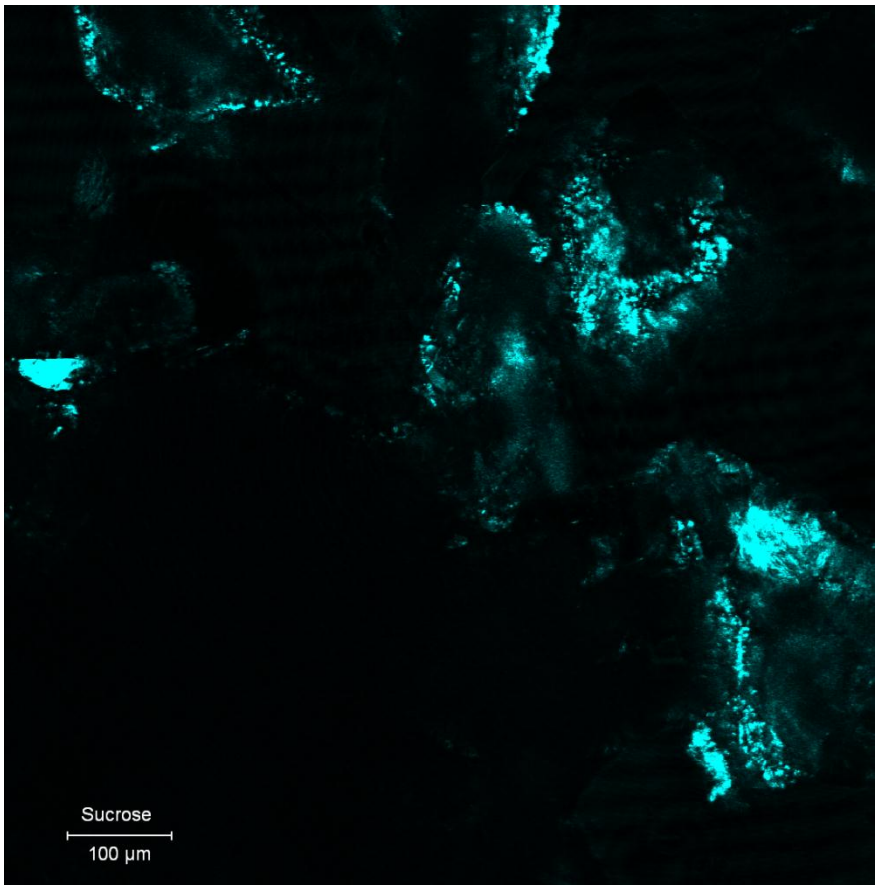



**Figure 6.14 A reflection image of sucrose taken using the CLSM at the Fonterra research centre. The image was provided by Elizabeth Nickless.**

The major difference between the techniques is the resolution. CLSM provides much better resolution than Raman microscopy. CLSM gives images where there are clearly defined borders between components with few regions of overlap, illustrated by the z-series' on the supplemental DVD. In the spherical cavities in the protein or AMF, there is a spherical region of AMF or fast green (intended to indicate the presence of RC), respectively. The model systems analysed using Raman microscopy also showed such cavities in component A being filled by component B, yet relatively large amounts of overlap were very common. The reason for this is the resolution of the instrument and the methodology. Resolution of CLSM (using the Zeiss LSM510) in the x/y plane is less than 0.33 µm, and can be as high as 0.8 µm in the z plane. It is dependent on the numerical aperture of the lens used, and the wavelength of the laser. Lateral resolution of the Raman microscope is 1.3 µm and was calculated considering the 50x objective (NA = 0.75) was used with the 785 nm laser. Vertical resolution is not

known at this stage and would require a series of signal tests performed on a silicon substrate at several depths using the particular objective and wavelength used. It is estimated to be greater than 5 μm based on work performed by McGoverin in 2007. It was expected to appear on the contour maps as overlap of the components along the borders of granules due to signal being received from both substances at once. This is observed in the AMF60SUC40 sample and a few of the AMF60RC40 contour maps. AMF60CS40 also has large amounts of overlap, but it is unclear whether this is due to resolution or the CS being dissolved in the AMF. As well as the limitations of the instrument, the methodology also has resolution limitations. This is due to the number of points at which spectra were taken over the analysed area. Pixel sizes were determined by the distance between the points ($\Delta X$ = 14.4 μm, $\Delta Y$ = 11.1 μm) and affected the amount of overlap when neighbouring pixels strongly indicate the presence of different components (and relatively small amounts of the same component). Since there is no data between the two points it is assumed that there is a gradient in intensity between the points and that mixing occurs. More points can be added to the grid hence creating much smaller pixels and thereby buffering the effect of this assumption. This was performed once on the AMF60SUC40 sample, shown in **figure 6.11**.However, inserting more measurement points did not noticeably decrease the amount of overlap between components. This suggests that the cause of the overlap is not the x/y resolution but more likely to be due to the innate disadvantage of the Raman microscope having much lower z-resolution than CLSM.

## 6.8 Conclusions

The large number of contour maps created using Raman microscopy and the few images obtained using confocal laser scanning microscopy revealed advantages and disadvantages for both techniques. Consistency between the numerous contour maps (between 2D and 3D maps and between layers within the 3D maps) attested to the reliability of the technique to show the location of the components. Furthermore, using different peaks to create the maps suggests that there is flexibility as well as reliability within the technique. Flaws in the methodology include fluorescence from the slide used and the way in which sample was unevenly placed on the slide. Both of these problems were easily fixed by simply using quartz slides with a cover slip. However, this experiment illustrated further flexibility in the technique as the effect of fluorescence was remedied by changing the integration method used. Resolution was the

largest problem and appeared as component overlap. This occurred in all samples but was most prevalent in AMF60CS40. Other than in AMF60CS40, it was established that most overlap occurred around the edges of granules and was attributed to relatively poor lateral and vertical resolution of the instrument and resolution of the methodology (which could be improved by using more data points in the map). CLSM can be seen to naturally have much better resolution but is at a disadvantage when looking at carbohydrates and other substances which cannot be stained. Perhaps most importantly, it can be seen that Raman microscopy can obtain images of individual components (within model systems) and ascertain their relative locations without the need for prior preparation or alteration.

# Chapter 7 Conclusions and Future Work

Raman spectroscopy proved its utility in both the analysis of paint samples and dairy products. Paint samples were analysed using the Senterra confocal Raman microscope to determine the identity of the pigments. Like many studies previously undertaken by Clark, the pigments of many of the samples were successfully elucidated. Future applications of this technique are likely to become more common simply because of its ease of use and the utility of identifying substances which were used in certain pieces of artwork for conservation and authenticity purposes.

Spiking of milk powder with calcium carbonate is an effective method for introducing extra calcium into the diet of a populace. Using Raman spectroscopy and principal component analysis (PCA), milk powder was able to be differentiated from samples spiked with calcium carbonate. Results showed that when concentrations of calcium carbonate only differed by 0.5% (w/w), principal component analysis was not able to differentiate samples effectively. Partial least squares (PLS) and univariate analysis were effective at predicting the concentration of calcite within the mixtures using Raman spectroscopic data. However, limit of quantification (LOQ) values indicated that neither technique would be useful for quantifying calcite concentrations as low as those used for calcium fortification in retail milk powder samples.

Cheese components, fat, protein, cornstarch and sucrose were also tested using the FT-Raman spectrometer and principal component analysis to determine whether separation of mixtures based on component concentrations could be achieved. Upon confirming the ability of the technique to do so, the ability of confocal Raman microscopy to map the components within the sample was investigated. Samples such as cheese are quite complex, which can hinder the ability to map them. In order to avoid this problem, model cheese systems containing two of four of the main components (fat, protein, cornstarch and sucrose) were constructed to investigate the ability of Raman microscopy to map simple mixtures of cheese components. Both 2D and 3D mapping was achieved, with results showing that even with flaws in technique and instrumentation, each component could be effectively located in the mixtures.

Reliability was established through the flexibility in the methodology used and through repetition which was required for constructing 3D maps. Comparisons were drawn between confocal laser scanning microscopy and confocal Raman microscopy, establishing the pitfalls of each technique; which includes resolution for Raman microscopy and the inability to stain certain samples for CLSM. Future work for this project should involve mapping of more complex samples as overlap between peaks from different components was found to be a hindrance; a situation which is much more likely to occur in more complex systems. The ability to map using the sum of several characteristic peaks should also be tested, as this showed promise for negating the influence of peak overlap. Throughout this project, Raman spectroscopy has illustrated great potential for application in the analysis of dairy products.

# References


1. R. J. H. Clark, L. Burgio, R. R. Hark, *Journal of Raman Spectroscopy* **2009**, *40*, 2031.
2. J. Zieba-Palus, A. Michalska, A. Weselucha-Birczynska, *Journal of Molecular Structure* **2011**, *993*, 134.
3. C. Lopez, B. Camier, J. Gassi, *International Dairy Journal* **2007**, *17*, 235.
4. L. Curda, O. Kukackova, *Journal of Food Engineering* **2004**, *61*, 557.
5. K. G. Adamopoulos, A. M. Goula, H. J. Petropakis, *Journal of Food Composition and Analysis* **2001**, *14*, 431.
6. G. Turrell, J. Corset, *Raman Microscopy Developments and Applications*, first edition, Academic Press, San Diego, **1996**.
7. K. Nakamoto, J. R. Ferraro, *Introductory Raman Spectroscopy*, first edition, Academic Press, San Diego, **1994.**
8. M. Windbergs, M. Haaser, C. M. McGoverin, K. C. Gordon, P. Kleinebudde, C. Strachan, *Journal of Pharmaceutical Sciences* **2010**, *99*, 1464.
9. C. M. McGoverin, *Raman Spectroscopy of Complex Mixtures*, Thesis (PhD), University of Otago, **2008**.
10. N. S. Claxton, T. J. Fellers, M. W. Davidson, *Laser Scanning Confocal Microscopy*, Florida State University, http://www.olympusfluoview.com/theory/LSCMIntro.pdf.
11. Y. Ozaki, *Journal of Adhesion Science and Technology* **2011**, *25*, 723.
12. S. R. Pygall, J. Whetstone, P. Timmins, C. D. Melia, *Advanced Drug Delivery Reviews* **2007**, *59*, 1434.
13. M. D. T. Clark, H. Wansbrough, S. Lipsham, *Paint and Pigments*, New Zealand Institute of Chemistry, http://nzic.org.nz/ChemProcesses/polymers/10D.pdf.
14. R. J. Gettens, G. L. Stout, *Painting Materials: A short Encyclopaedia*, Dover edition, Dover Publishers, Inc., New York, **1966.**
15. R. J. H. Clark, *Chemistry in New Zealand* **2011**, *75*, 13.
16. A. M. Correia, R. J. H. Clark, M. I. M. Ribeiro, M. L. T. S. Duarte, *Journal of Raman Spectrocopy* **2007**, *38*, 1390.
17. R. J. H. Clark, S. Mirabaud, *Journal of Raman Spectroscopy* **2006**, *37*, 235.

18. R. J. H. Clark, P. J. Gibbs, K. R. Seddon, N. M. Brovenko, Y. A. Petrosyan, *Journal of Raman Spectroscopy* **1997**, *28*, 91.
19. R. J. H. Clark, J. van der Weerd, *Journal of Raman Spectroscopy* **2004**, *35*, 279.
20. T. D. Chaplin, A. Jurado-Lopez, R. J. H. Clark, D. R. Beech, *Journal of Raman Spectroscopy* **2004**, *35*, 600.
21. T. D. Chaplin, R. J. H. Clark, D. A. Scott, *Journal of Raman Spectroscopy* **2006**, *37*, 223.
22. O. R. Fennema (editor), *Food Chemistry*, third edition, Marcel Dekker, Inc., New York, **1996.**
23. H. Soyeurt, P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, N. Gengler, *Journal of Dairy Science* **2006**, *89*, 3690.
24. P. Fox, P. L. H. McSweeney, T. M. Cogan, T. P. Guinee, *Fundamentals of Cheese Science*, first edition, Aspen Publications, Gaithersburg, **2000**.
25. M. Caric, *Concentrated and Dried Dairy Products*, first edition, VCH Publishers, Inc., New York, **1994**.
26. World Health Organization, *Toxicological and Health aspects of Melamine and Cyanuric Acid*, Report of a WHO Expert Meeting in Collaboration with FAO, Geneva, **2009**.
27. B. Kim, L. B. Perkins, R. J. Bushway, S. Nesbit, T. Fan, R. Sheridan, V. Greene, *Journal of AOAC International* **2008**, *91*, 408.
28. Y. Cheng, Y. Dong, J. Wu, X. Yang, H. Bai, H. Zheng, D. Ren, Y. Zou, M. Li, *Journal of Food Composition and Analysis* **2010**, *23*, 199.
29. K. R. Beebe, R. J. Pell, M. B. Seasholtz, *Chemometrics: A Practical Guide*, first edition, John Wiley and Sons, New York, **1998**.
30. H. M. Barker, *Nutrition and Dietetics for Health Care*, tenth edition, Elsevier Health Sciences, Philadelphia, **2002**.
31. R. M. El-Abassy, P. J. Eravuchira, P. Donfack, B. von der Kammer, A. Materny, *Vibrational Spectroscopy* **2011**, *56*, 3.
32. C. M. McGoverin, A. S. S. Clark, S. E. Holroyd, K. C. Gordon, *Analytical Methods* **2009**, *1*, 29.
33. G. R. Flåten, A. D. Walmsley, *Analyst* **2003**, *128*, 935.

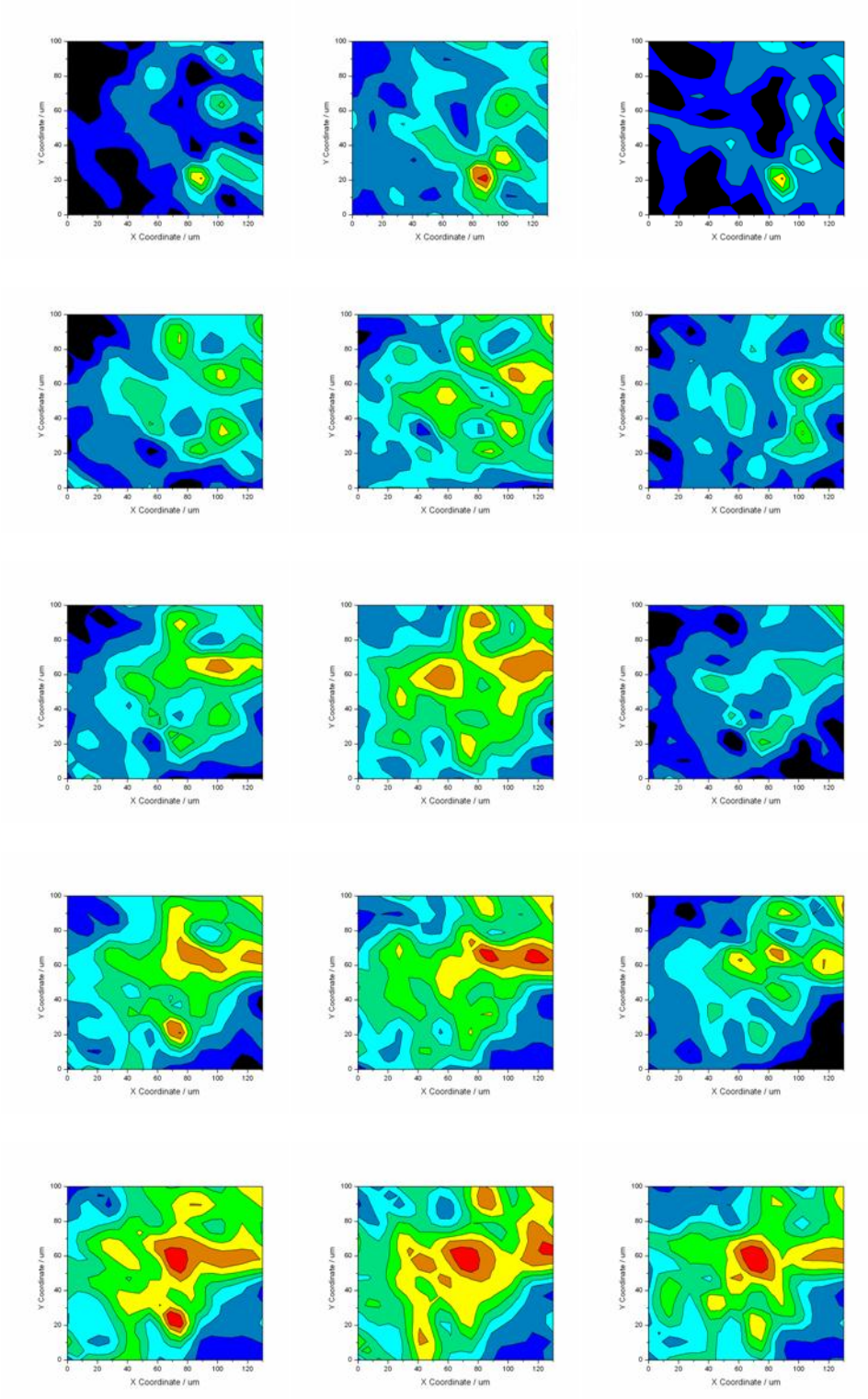
34. S. Gallier, K. C. Gordon, R. Jiménez-Flores, D. W. Everett, *International Dairy Journal* **2011**, *21*, 402.
35. E. C. Y. Li-Chan, *Trends in Food Science and Technology* **1996**, *7*, 361.
36. C. M. McGoverin, A. S. S. Clark, S. E. Holroyd, K. C. Gordon, *Analytica Chimica Acta* **2010**, *673*, 26.
37. J. E. Romanchik-Cerpovics, R. J. McKemie, *Journal of the American Dietetic Association* **2007**, *107*, 506.
38. M. R. Alameida, K. de S. Oliveira, R. Stephani, L. F. C. de Oliveira, *Journal of Raman Spectrocopy* **2011**, *42*, 1548.
39. J. Fontecha, J. Bellanato, M. Juarez, *Journal of Dairy Science* **1993**, *76*, 3303.
40. I. M. Bell, R. J. H. Clark, P. J. Gibbs, *Raman Spectroscopic Libraries*, http://www.chem.ucl.ac.uk/resources/raman/index.html.
41. I. M. Bell, R. J. H. Clark, P. J. Gibbs, *Spectrochimica Acta Part A* **1997**, *53*, 2159.
42. T. Lana-Villarreal, J. M. Perez, R. Gomez, *C. R. Chimie* **2006**, *9*, 806.
43. M. Gniadecka, H. C. Wulf, O. F. Nielson, D. H. Christensen, J. Hercogova, *Photochemistry and Photobiology* **1997**, *66*, 418.
44. R. Aparicio, V. Baeten, P. Hourant, M. Morales, *Journal of Agricultural Food Chemistry* **1998,** *46*, 2638.
45. D. D. Miller, *Advances in Food and Nutrition Research* **1989**, *33*, 103.
46. H. K. Vyas, P. S. Tong, *Journal of Dairy Science* **2004**, *87*, 1177.
47. A. Candolfi, R. De Maesschalck, D. Jouan-Rimbaud, P. A. Hailey, D. L. Massart, *Journal of Pharmaceutical and Biomedical Analysis* **1999**, *21*, 115.
48. K. H. Esbensen, *Multivariate Data Analysis in Practice*, fifth edition, CAMO Process AS, Oslo, **2002**.
49. K. G. Adamopoulos, A. M. Goula, H. J. Petropakis, *Journal of Food Composition and Analysis* **2001**, *14*, 431.
50. L. Curda, O. Kukackova, *Journal of Food Engineering* **2004**, *61*, 557.
51. D. M. Byler, H. M. Farrell, JR., H. Susi, *Journal of Dairy Science* **1988**, *71*, 2622.
52. L. Silveira Jr., L. M. Moreira, V. G. B. Conceição, H. L. Casalechi, I. S. Muñoz, F. F. da Silva, M. A. S. R. Silva, R. A. de Souza, M. T. T. Pacheco, *Spectroscopy* **2009**, *23*, 217.

53. M. R. Almeida, R. S. Alves, L. B. L. R. Nascimbem, R. Stephani, R. J. Poppi, L. F. C. de Oliveira, *Analytical and Bioanalytical Chemistry* **2010**, *397*, 2693.
54. C. Lopez, V. Briard-Bion, B. Camier, J.-Y. Gassi, *Journal of Dairy Science* **2006**, *89*, 2894.
55. J. R. Schoonover, S. L. Zhang, C. T. Johnston, *Journal of Raman Spectroscopy* **2003**, *34*, 404.
56. R. Tauler, B. Kowalski, S. Fleming, *Analytical Chemistry* **1993**, *65*, 2040.

# Appendix A Stage Movement in the Z-Axis

During the time in which the 3D contour maps were being created, there was some uncertainty regarding the accuracy of the stage movement in the z-direction. It wasn't until the Bruker technician came to fix an unrelated problem with the microscope and to run several diagnostic tests that the z-movement inaccuracy became apparent. It was found that when the microscope was programmed to move the stage a specific distance in the z-axis, it would only move 40% of the desired distance instead. This was found to be consistent and precisely 40% of the instructed movement. This problem has not been fixed yet, though has been avoided thus far by multiplying the desired stage movement by 2.5 and programming the stage to move the resulting distance instead. The 3D contour maps were created before the technician arrived; therefore it was assumed that prior to this occasion, the stage movement was still limited to 40% of the programmed distance. The reported 3.2 µm distance between adjacent contour maps is an adjusted value. The distance that the stage was programmed to move (and hence was originally thought to be the distance moved by the stage) was in fact 8 µm.

# Appendix B Contour Maps from 3D Mapping



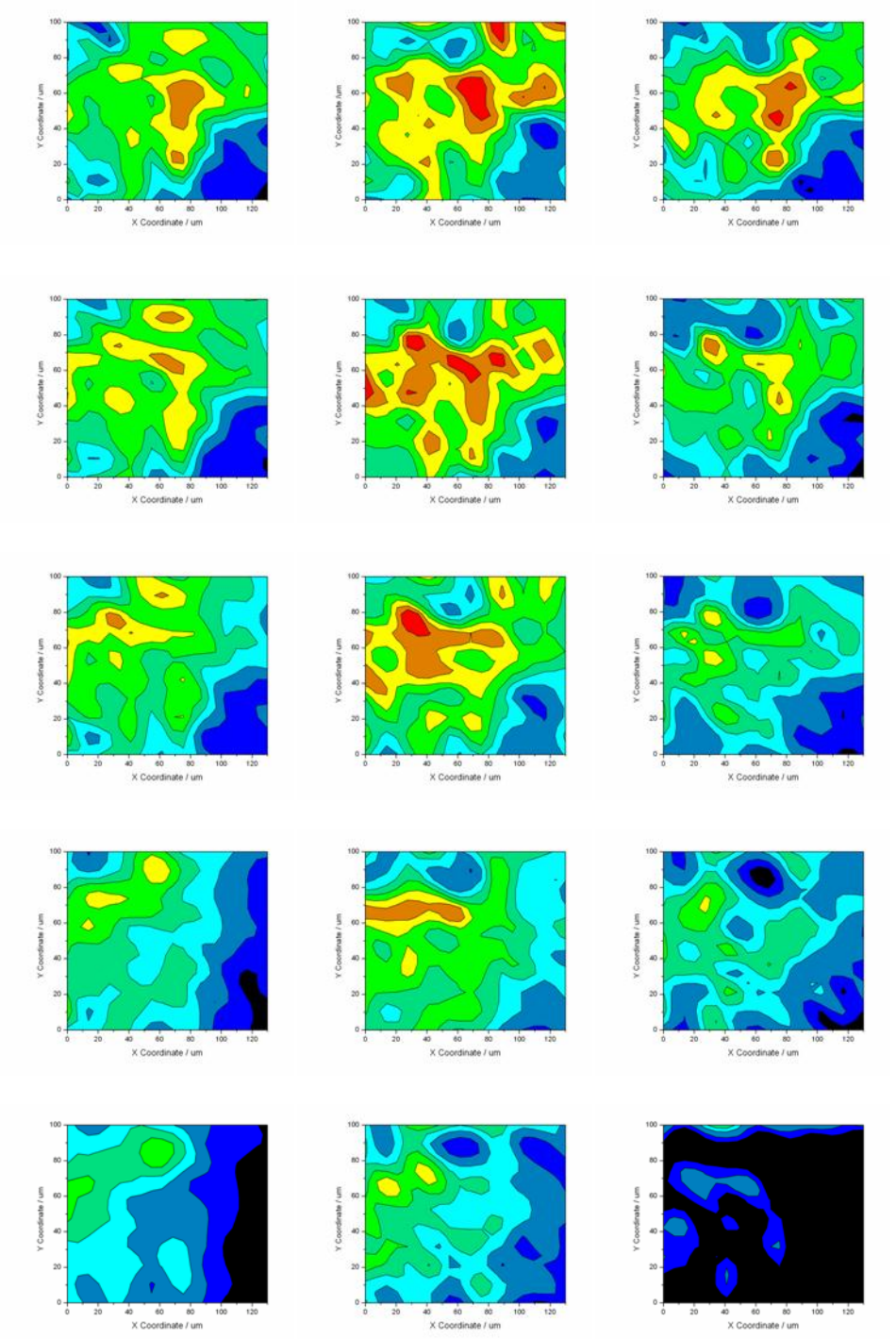
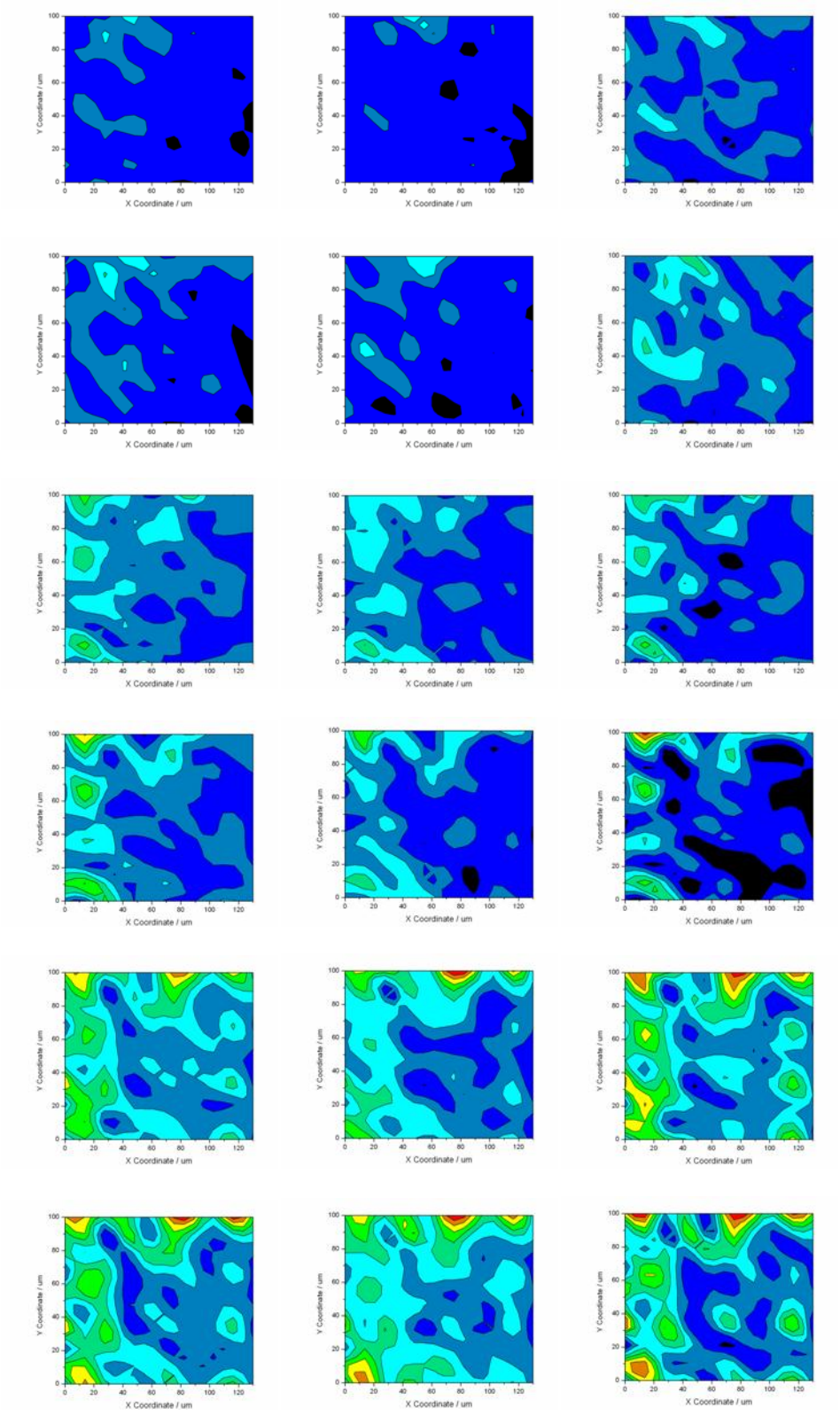



**Figure 8.1 Three columns of contour maps for AMF60SUC40 mixture. Each column of 10 contour maps has been created using a different peak with the top contour map of each column being the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer. All three peaks used are characteristic of the presence of AMF. The columns consist of: contour maps created using the 1440 $cm^{-1}$ peak (left), maps created using the 1655 $cm^{-1}$ peak (middle), and maps created using the 1744 $cm^{-1}$ peak (right).**


Y Coordinate / um
X Coordinate / um
0
20
40
60
80
100
120

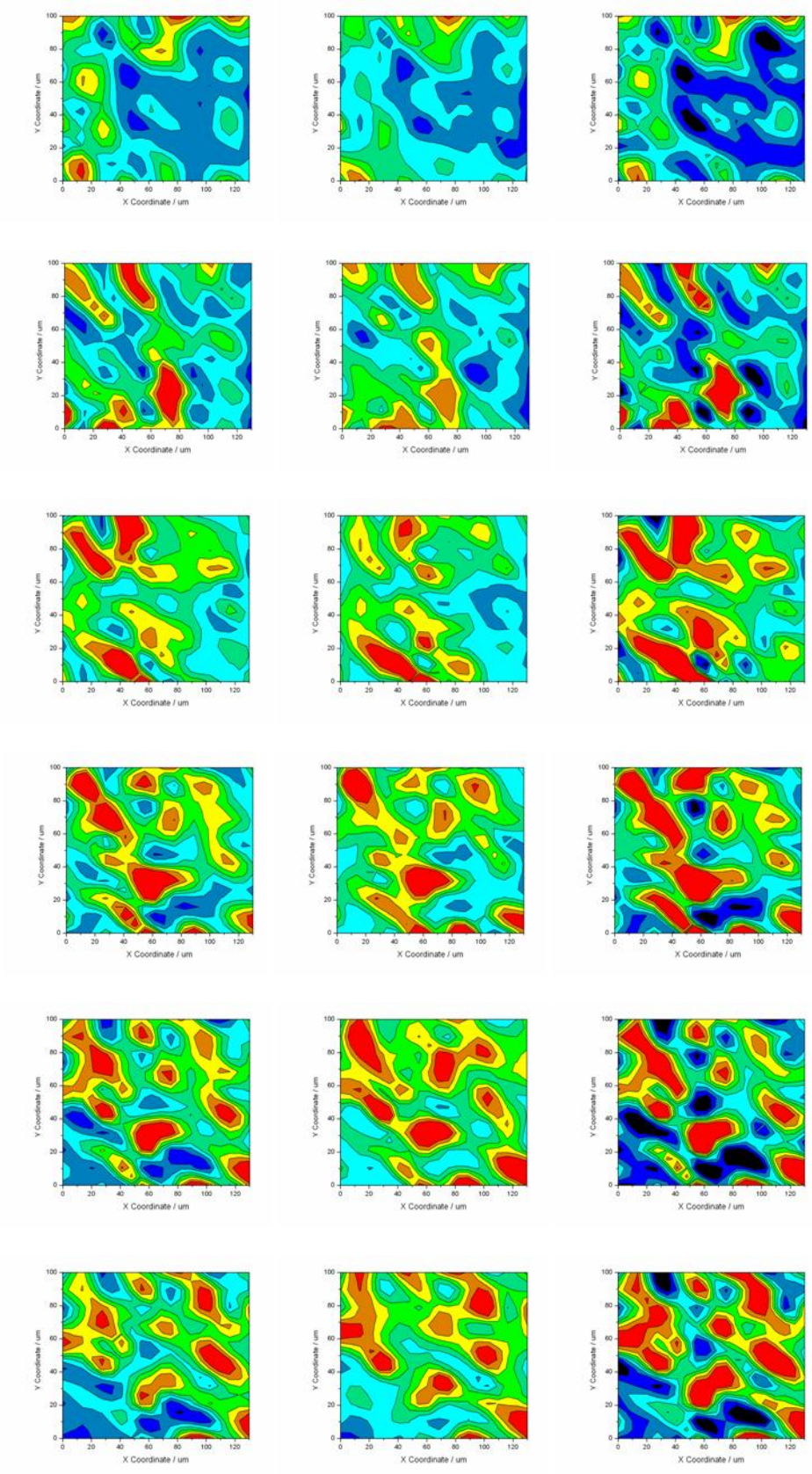
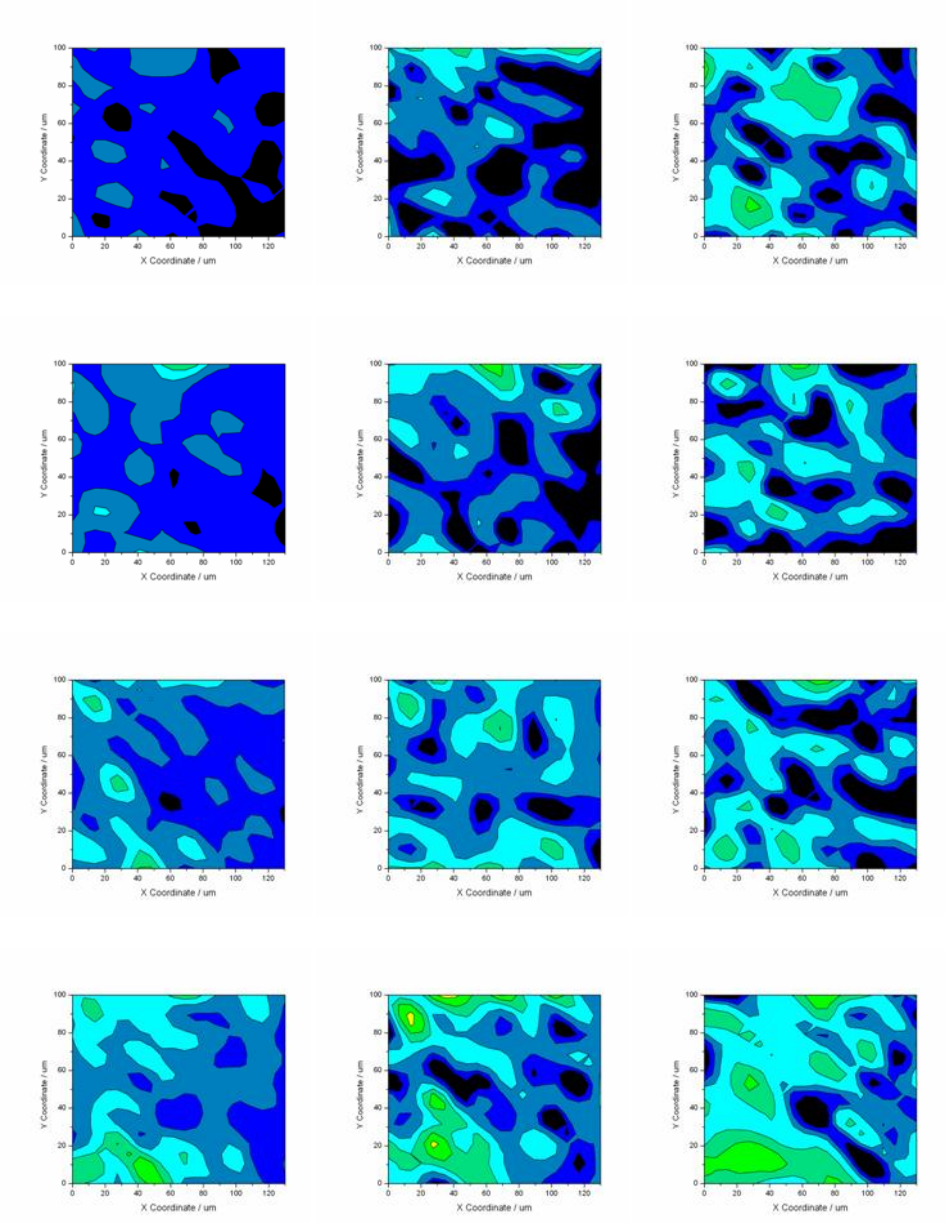
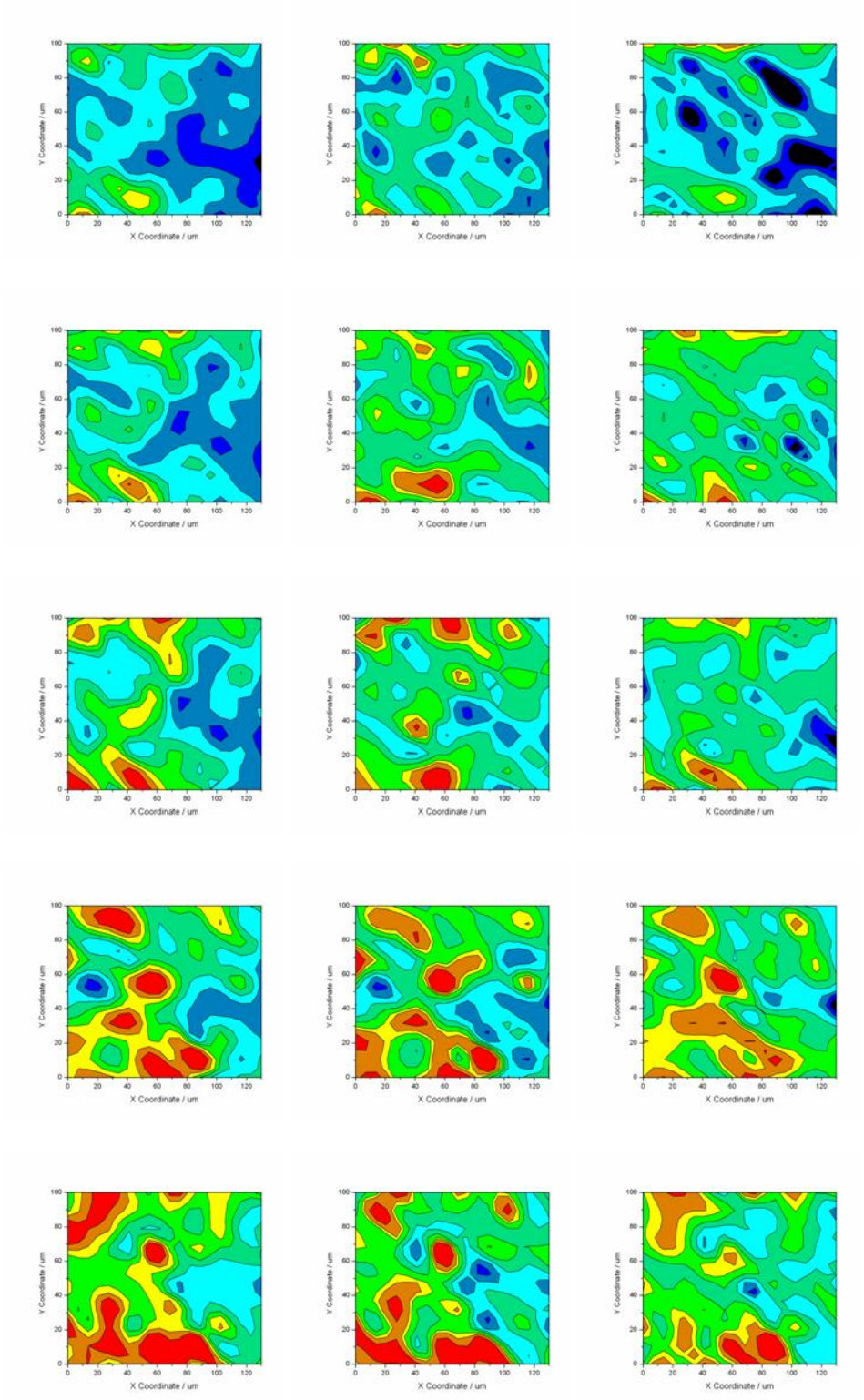

Y Coordinate / um
X Coordinate / um
0
20
40
60
80
100
120

**Figure 8.2 Three columns of contour maps for AMF60CS40 mixture. Each column of 12 contour maps has been created using a different peak with the top contour map of each column being the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer. All three peaks used are characteristic of the presence of CS. The columns consist of: contour maps created using the 477 $cm^{-1}$ peak (left), maps created using the 865 $cm^{-1}$ peak (middle), and maps created using the 940 $cm^{-1}$ peak (right).**




Y Coordinate / um
X Coordinate / um
0
20
40
60
80
100
120

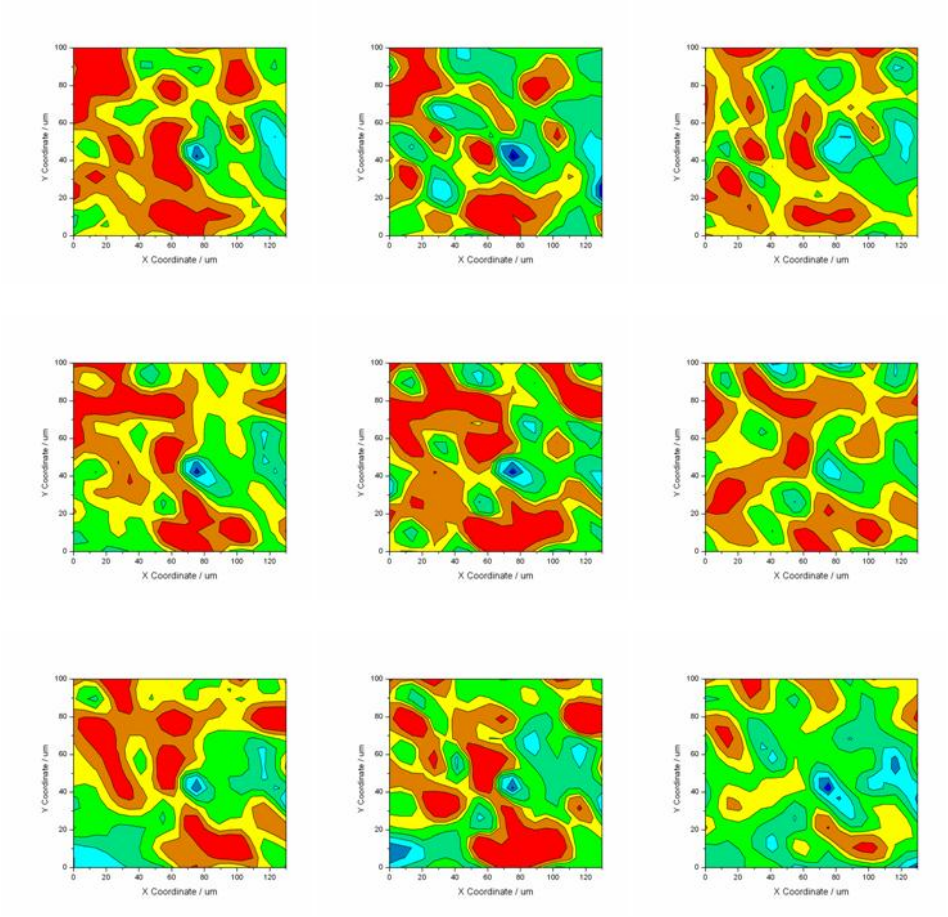
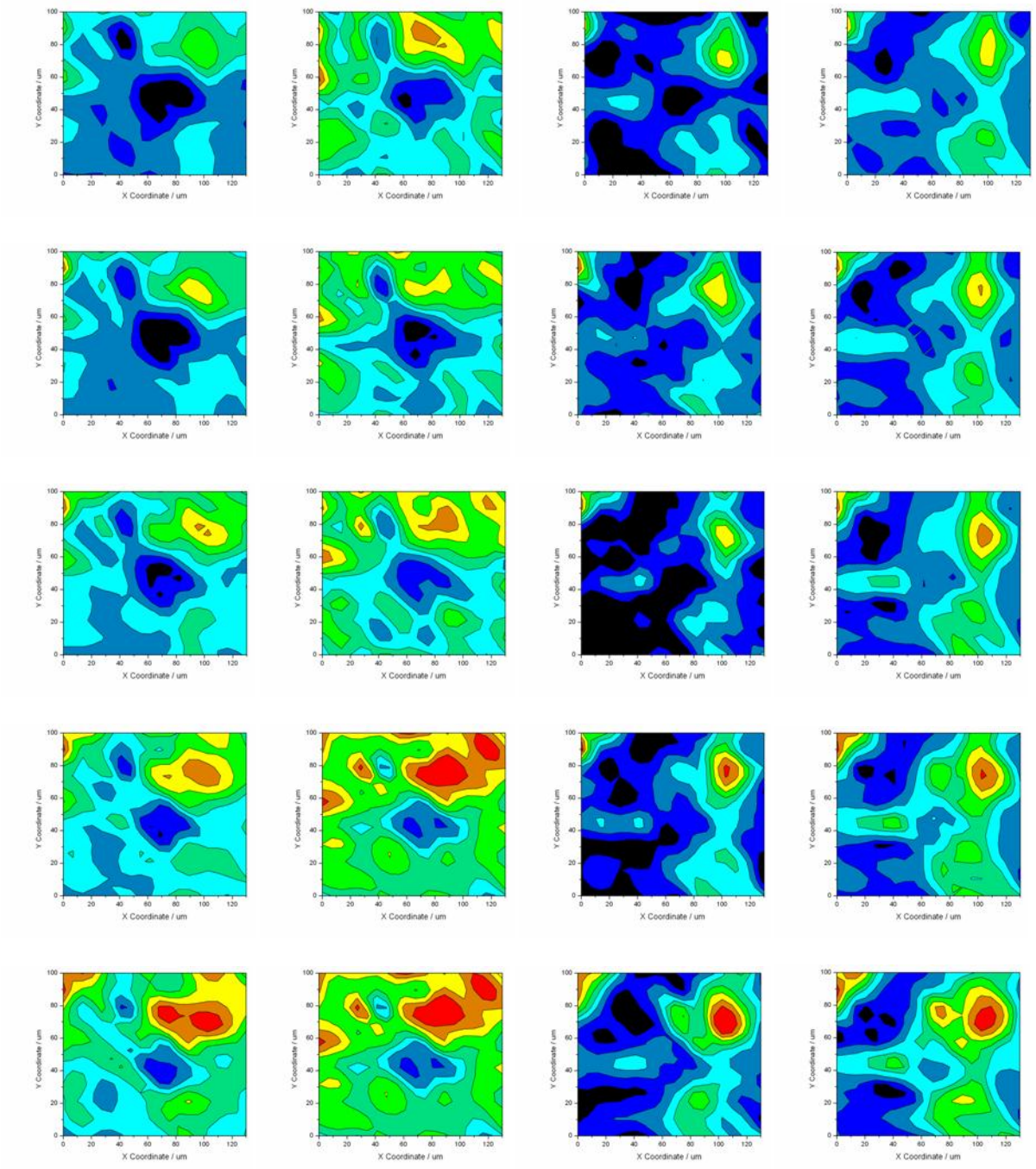



**Figure 8.3 Three columns of contour maps for AMF60CS40 mixture. Each column of 12 contour maps has been created using a different peak with the top contour map of each column being the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer. All three peaks used are characteristic of the presence of AMF. The columns consist of: contour maps created using the 1440 $cm^{-1}$ peak (left), maps created using the 1655 $cm^{-1}$ peak (middle), and maps created using the 1744 $cm^{-1}$ peak (right).**


Y Coordinate / um
X Coordinate / um
0 20 40 60 80 100 120

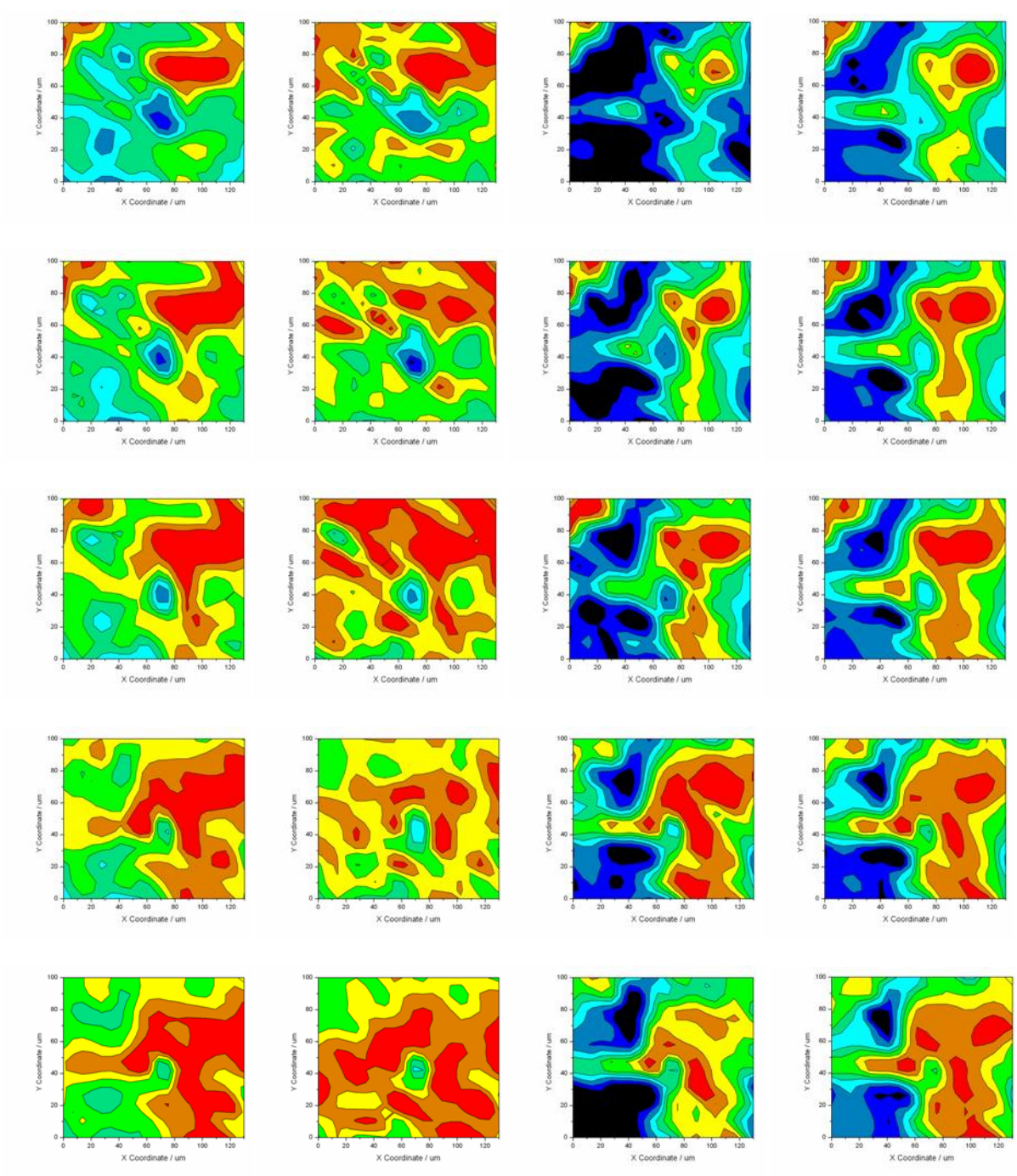



**Figure 8.4 Four columns of contour maps for AMF60RC40 mixture. Each column of 10 contour maps has been created using a different peak with the top contour map of each column being the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer. All three peaks used are characteristic of the presence of AMF. The columns consist of: contour maps created using the 1440 $cm^{-1}$ peak (left), maps created using the 1655 $cm^{-1}$ peak (middle left), maps created using the 1744 $cm^{-1}$ peak (middle right), and maps created using the 1300 $cm^{-1}$ peak (right).**

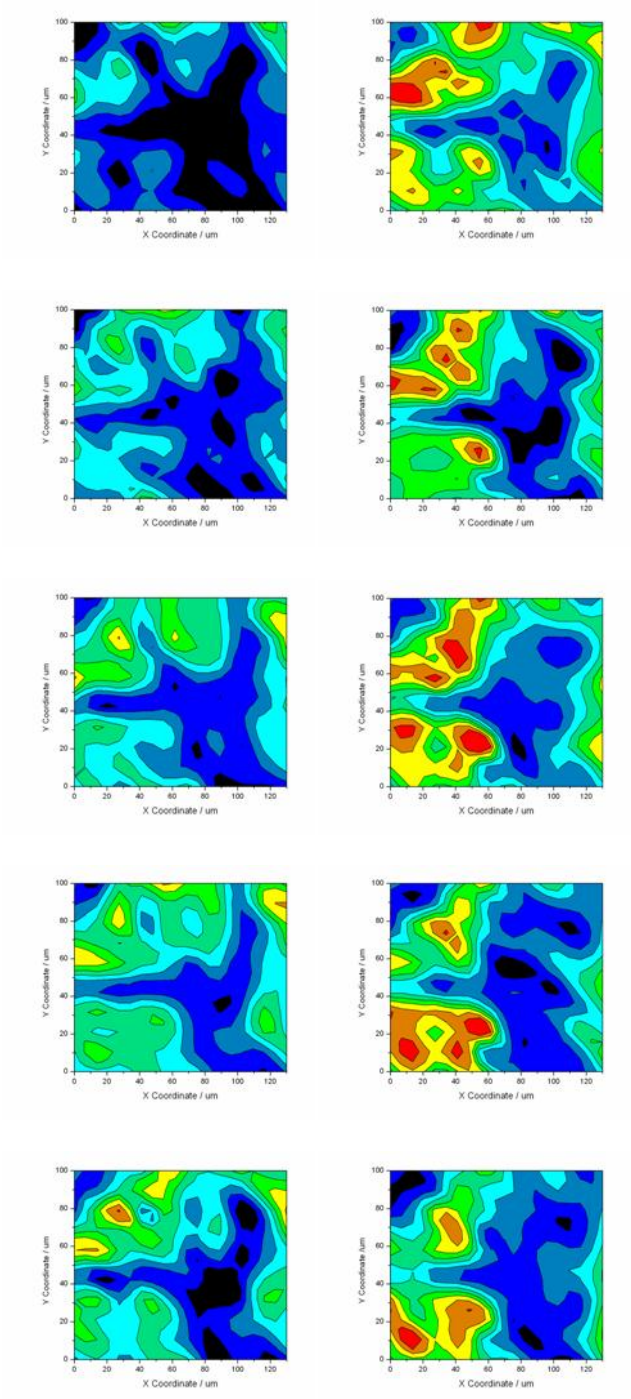



**Figure 8.5 Two columns of contour maps for AMF60RC40 mixture. Layers 1 to 5 (left) and layers 6-10 (right). The top contour map of each column is the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer (the topmost layer is the contour map at the bottom of the right-hand column). All contour maps belong to the RC peak at 1003 $cm^{-1}$.**

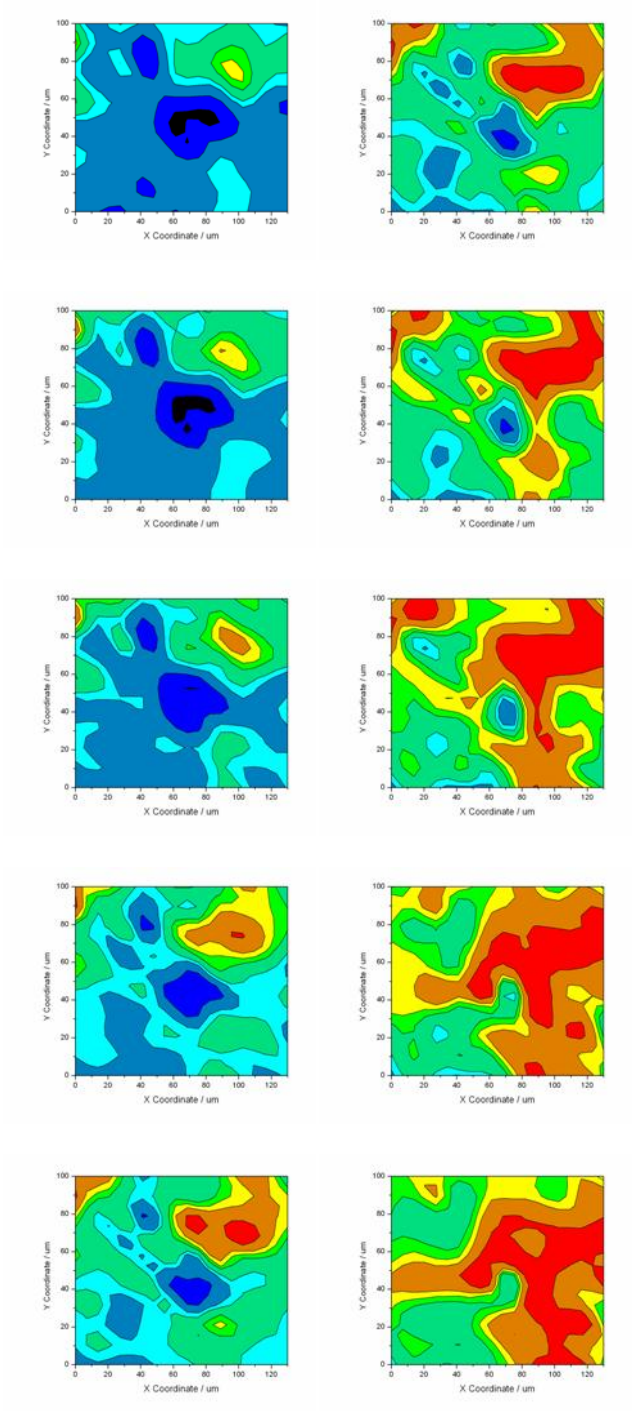
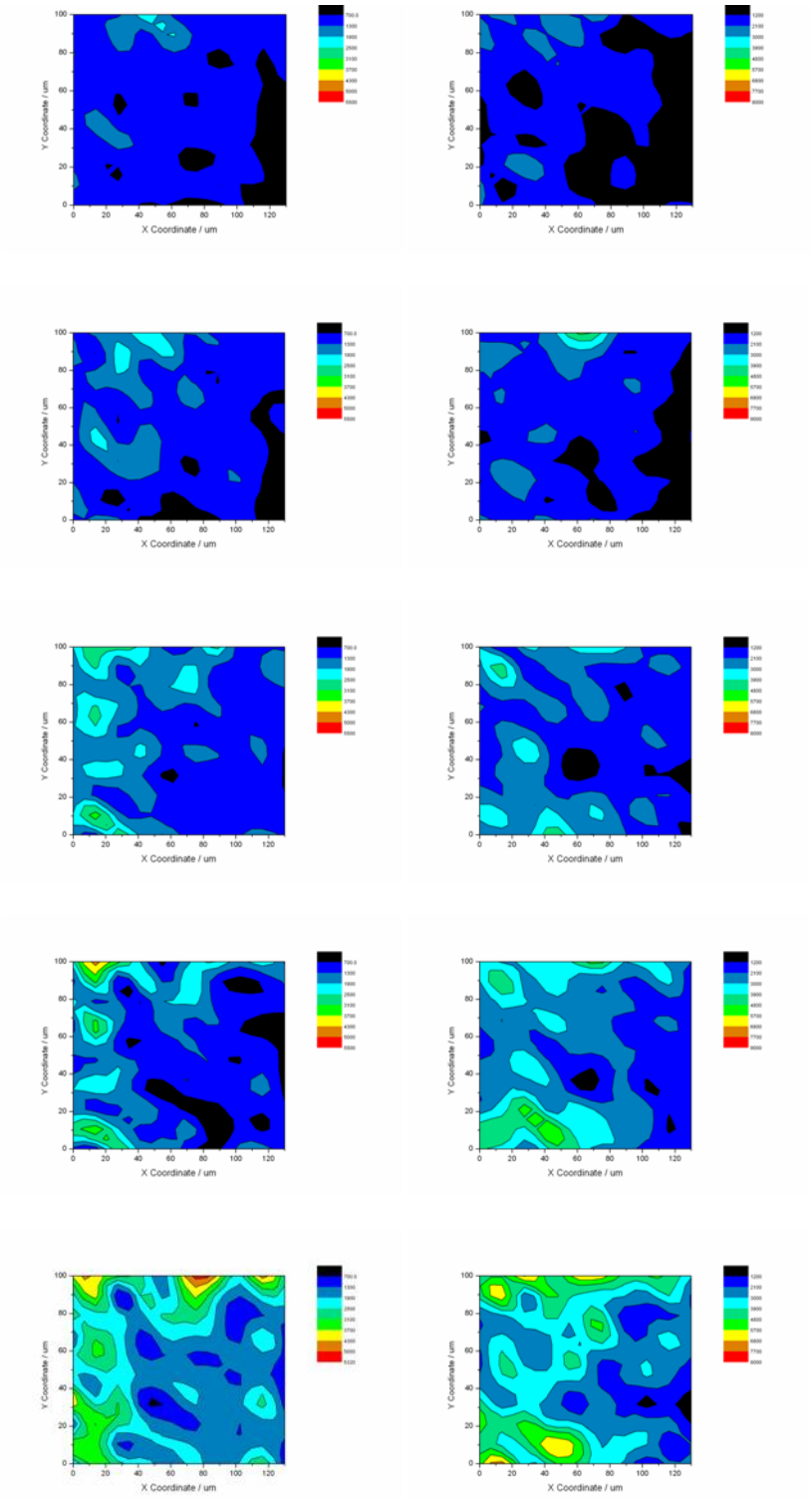



**Figure 8.6 Two columns of contour maps for AMF60RC40 mixture. Layers 1 to 5 (left) and layers 6-10 (right). These contour maps were formed by adding the integration values for three AMF peaks together. The AMF peaks used include the 1440, 1655 and 1744 $cm^{-1}$ peaks. The top contour map of each column is the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer (the topmost layer is the contour map at the bottom of the right-hand column).**


Y Coordinate / um
X Coordinate / um

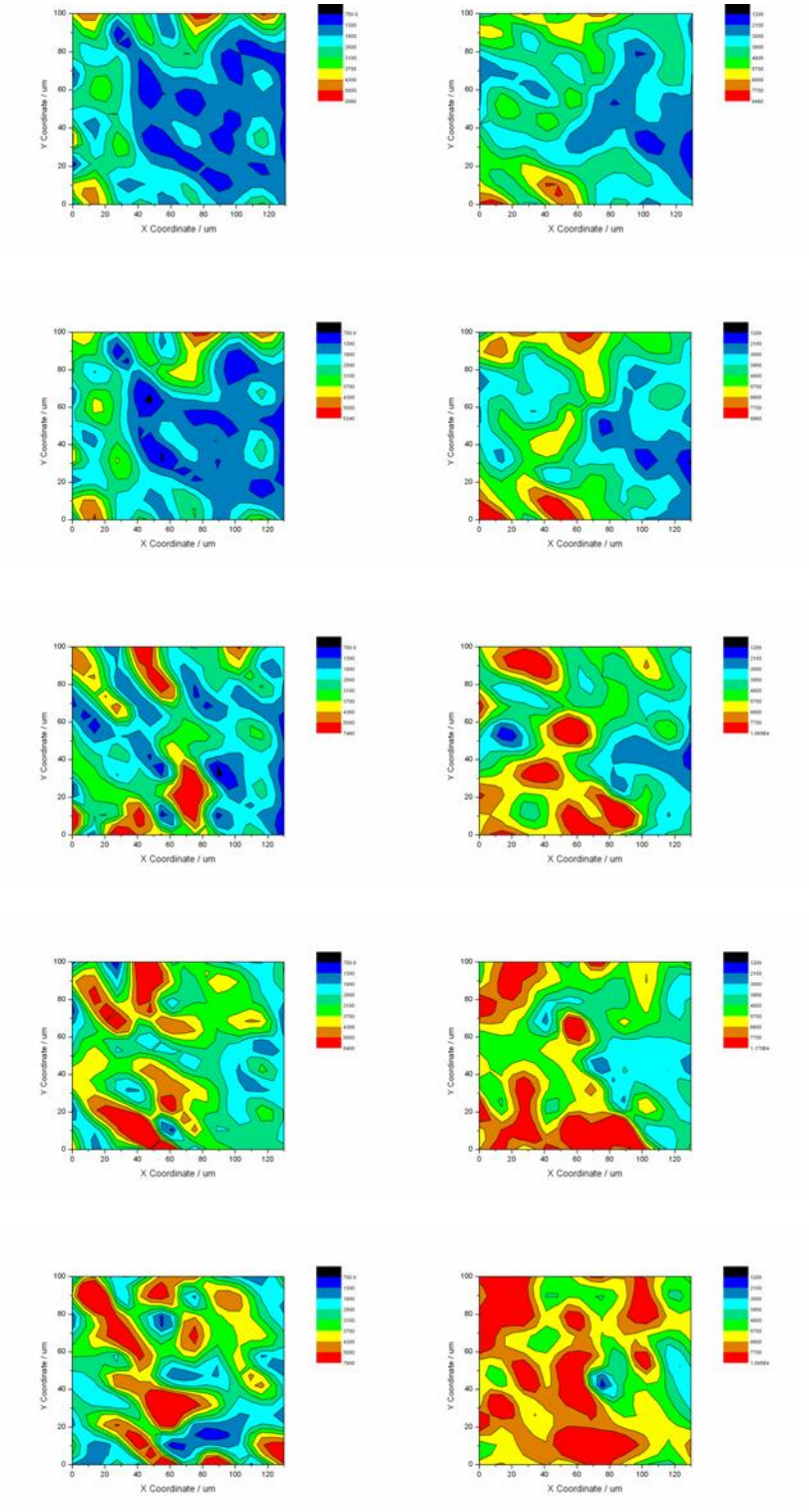
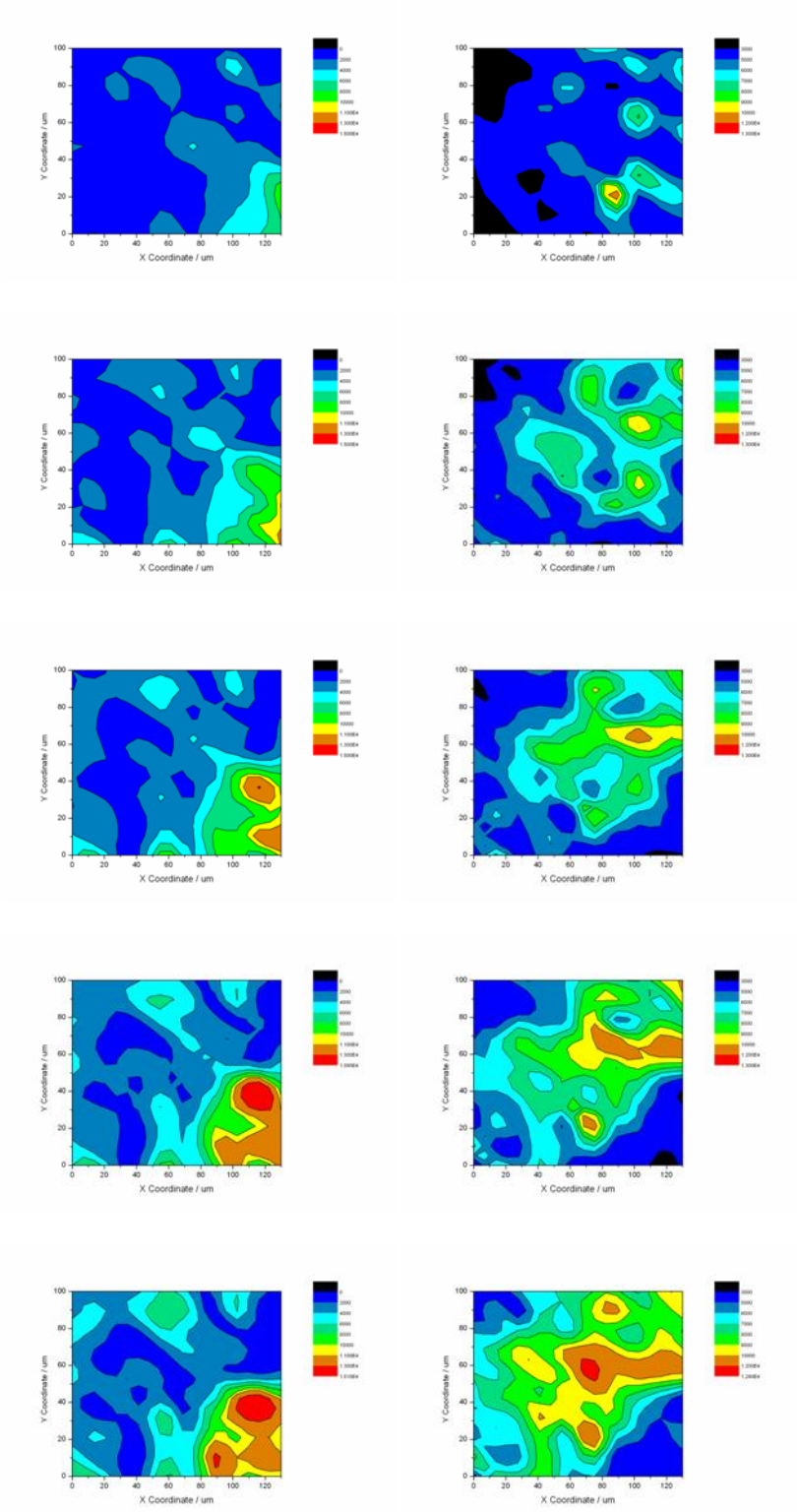



**Figure 8.7 Two columns of contour maps for AMF60CS40 mixture. The contour maps in the left column were created from the sum of the 477, 865 and 940 $cm^{-1}$ CS peaks. The contour maps in the right column were created from the sum of the 1440, 1655 and 1744 $cm^{-1}$ AMF peaks. The top contour map of each column is the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer.**


Y Coordinate / um
X Coordinate / um

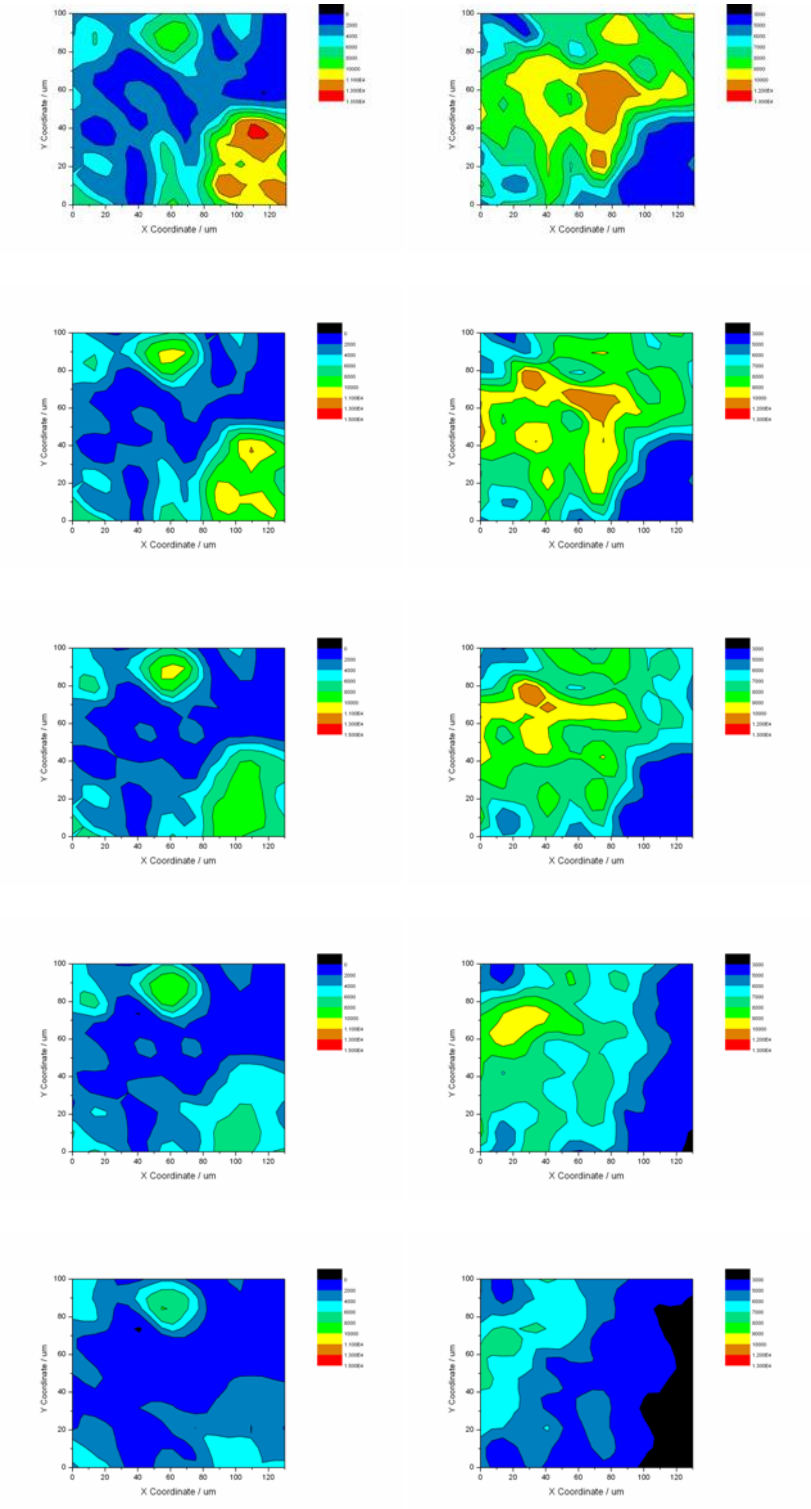



**Figure 8.8 Two columns of contour maps for AMF60SUC40 mixture. The contour maps in the left column were created from the sum of the 539, 641 and 849 $cm^{-1}$ SUC peaks. The contour maps in the right column were created from the sum of the 1440, 1655 and 1744 $cm^{-1}$ AMF peaks. The top contour map of each column is the lowest layer (z coordinate) and each row is 3.2 µm above the previous layer.**