Investigations of Atherosclerotic Coronary Arteries by Confocal Raman Microspectroscopy

by

Henry C.Y. Chang

B.Sc.(Hon), The University of British Columbia, 2000

A Thesis Submitted in Partial Fulfillment of

the Requirements for the Degree of

Master of Science

in

The Faculty of Graduate Studies

Department of Chemistry

We accept this thesis as conforming to the required standard

The University of British Columbia

January 2003

In presenting this thesis in partial fulfilment of the requirements for an advanced degree at the University of British Columbia, I agree that the Library shall make it freely available for reference and study. I further agree that permission for extensive copying of this thesis for scholarly purposes may be granted by the head of my department or by his or her representatives. It is understood that copying or publication of this thesis for financial gain shall not be allowed without my written permission.

Department of CHEMISTRY

The University of British Columbia
Vancouver, Canada

Date Jan 3rd, 2003

ABSTRACT

Atherosclerosis is a vascular disease, initiated by damage to the endothelial cells, and characterized by the deposition of plaque, which composes of mainly lipids and cholesterol, in the inner arterial wall that leads to progressive narrowing and hardening of the artery. The assessment of plaque status for the prediction of disease progression requires the examination of plaque's composition and structure rather than its size or gross morphology. Recent studies have shown Raman spectroscopy to be well suited for diagnosis of tissues because of its ability to identify and quantify biochemical molecules and structural motifs within intact cells. The ultimate utility of Raman technique may lie in its potential for the real-time assessment of atherosclerosis *in vivo* using fiber optic probes. This project employed confocal Raman microspectroscopy for the study of normal and diseased human coronary artery specimens *in situ*.

A new Raman microscopy system for tissue spectroscopy has been set-up and characterized. Raman spectra were collected at various locations of the arterial wall, either along the radius covering its different layers, or concentrically outward from the lumen. Numerous algorithms were developed and tested for baseline removal, normalization and smoothing of Raman spectra obtained. Key analysis involving component analysis and generalized two-dimensional (2D) correlation followed. A model for chemical component analysis of eighteen major biochemicals found in the tissue/plaque was developed based on the application of singular value decomposition (SVD). Quantitative information on the relative amount of individual biomolecule was provided by the component analysis in which each Raman artery spectrum was modeled

to a set of eighteen biochemical basis spectra. Generalized 2D Raman correlation spectroscopy was employed to identify and classify variations in spectral features occurring at different locations of the artery transverse section. The 2D correlation results have demonstrated the ability of 2D correlation to detect subtle spectral changes happening within a tiny area of the arterial wall, as well as to provide qualitative information about the order of appearance of individual constituent that complemented the chemical composition results. Possible future improvements and continuation of the project were proposed in the final chapter of this thesis as well.

TABLE OF CONTENTS

List of Figures

List of Abbreviations

1D	One-Dimensional
2D	Two-Dimensional
3D	Three-Dimensional
CaF_2	Calcium Fluoride
CAD	Coronary Artery Disease
CCD	Charge Coupled Device
CHD	Coronary Heart Disease
CRM	Confocal Raman Microspectrometer
CT	Computed Tomography
CVD	Cardiovascular Disease
CW	Continuous Wave
EBCT	Electron Beam Computed Tomography
FT	Fourier Transform
FWHM	Full Width at Half-Maximum
GRAMS/32®	Graphic Relational Array Management System
HLBF	Holographic Laser Bandpass Filter
HNF	Holographic Notch Filter
IHD	Ischemic Heart Disease
IR	Infrared
IVUS	Intravascular Ultrasound
MRI	Magnetic Resonance Imaging

$Nd:YVO_4$	Neodymium Vanadate
NIR	Near Infrared
NMR	Nuclear Magnetic Resonance
OCT	Optimal Cutting Temperature
PCA	Principal Component Analysis
PET	Positron Emission Tomography
PIXE	Particle Induced X-ray Emission
RM	Raman Microspectroscopy
S/N	Signal-To-Noise
SDEV	Standard Deviation
SMCs	Smooth Muscle Cells
SNR	Signal-To-Noise Ratio
SVD	Singular Value Decomposition
$TiAl_2O_3$	Titanium:Sapphire
UV	Ultraviolet
UVRRS	Ultraviolet Resonance Raman Spectroscopy
w/w	Weight percent (weight of solute / weight of solution)
WHO	World Health Organization
WiRE™	Window® based Raman Environment

Acknowledgements

First and foremost, I would like to thank my supervisors Dr. Michael W. Blades (Department of Chemistry and The Laboratory of Molecular Biophysics, U.B.C.) and Dr. Robin F.B. Turner (Department of Electrical and Computer Engineering and The Biotechnology Laboratory, U.B.C.), who allowed me complete freedom to pursue research in their laboratories, and provided invaluable advice and knowledge at times when needed, enabling the completion of this project. As well, I wish to thank Dr. Georg H. Schulze for his assistance and guidance in numerous aspects throughout the past two years. I also wish to thank Dr. Shane L. Greek for carrying out preliminary work that helped to define the objectives and scope of the project. I am very grateful to many members of the group, Christopher Barbosa, Ken Wright, Shengyong Lu, August Specht, Adam Bass and Chris Sherwood for their invaluable assistance and those memorable moments in the laboratories. Furthermore, special thanks to Mark Chen and Arnel Lim, two undergraduate summer students who contributed significantly towards the successful completion of this project.

I am indebted to the critical guidance and help from Dr. Michael Allard (Department of Pathology, St. Paul's Hospital), Mr. Albert Lee (Department of Pathology, St. Paul's Hospital), and those technicians from St. Paul's Hospital for providing and preparing artery tissue specimens. In addition, I gratefully acknowledge members of the Mechanical and Electrical Services Shops for their assistance.

The project was financially supported by grants from the University of British Columbia (UBC), the Natural Sciences and Engineering Research Council of Canada (NSERC), the Canadian Foundation for Innovation (CFI), BC Knowledge Development Fund, and the Laboratory of Molecular Biophysics (LMB) of UBC.

Finally yet importantly, I would like to thank my parents, my younger brother and my younger sister for the love, support and patience during the course of the study and the preparation of this thesis.

Chapter 1

Introduction

1.1 Overview: Atherosclerosis

The word "atherosclerosis" comes from two Greek terms: i) athero, which means paste or gruel, and ii) sclerosis, which means hardness. Atherosclerosis involves the buildup of a plaque in the inner arterial wall, initiated by a damage (called a lesion) to the endothelial cells there, followed by deposition of phospholipids, cholesterol, cellular debris, platelets, calcium salts and other substances in this inner lining of the blood vessel [1,2]. Varying proportions of these components occur in different plaques, thus giving rise to different forms of plaque. The progression of atherosclersis will be discussed in more detail in the following section. In atherosclersis, a plaque may become large enough to thicken the endothelium significantly. The artery diameter shrinks and blood flow decreases, reducing the oxygen supply.

Sometimes a plaque can also become unstable and rupture. When a ruptured plaque causes a blood clot to form (called thrombus), blood flow in the artery can be totally blocked. The blood clot may even break off and travel to other blood vessels of the body. A clot can block a blood vessel that feeds the heart or a blood vessel that feeds the brain, and it results in a heart attack or a stroke respectively [1]. The term ischemia means the condition when the flow of blood, and therefore oxygen, to a part of the body is restricted. Ischemic heart disease (IHD) refers to reduced or stoppage of blood flow and oxygen to the heart muscle caused by severely narrowed heart arteries. The disease is also called coronary artery disease (CAD) or coronary heart disease (CHD), and its underlying cause is mainly coronary arterial atherosclerosis [2]. It is known that the risk

of plaque stability is a function more of its composition than of its size [3,4]. Thus, assessment of the morphology and chemical compositional structure of plaques is of far greater diagnostic and prognostic value than measuring the degree of stenosis (or occlusion) of the artery.

The heart and the system of veins and arteries throughout the whole body are collectively known as the cardiovascular system. Cardiovascular disease (CVD) accounts for the deaths of 17 million people around the globe each year (one-third of global deaths), according to the estimation done by the World Health Organization (WHO) in 1999 [5,6]. During the year, CVD was reported as the leading cause of death in 31 of 35 countries reporting mortality statistics. Heart attack and stoke, both primarily atherosclerotic in origin, contribute to 7.3 million and 5.1 million deaths respectively. It was estimated that approximately 8 million Canadians and more than 61 million Americans have some sort of cardiovascular disease [7,8]. 600 million people worldwide with high blood pressure are at risk of CVD [9]. The aging global population will only serve to increase these numbers.

No reliable method can predict in advance which lesions will progress, regress, or rupture [10]. However, with the advent of new predictive technologies, there will be an opportunity for intervention based on early detection of disease. Optical systems provide access to many organs in the body with the use of flexible optical fibers to excite and collect the spectroscopic signals [11,12]. In addition, optical technologies offer real-time assessment of biochemical and morphologic tissue composition, and provide an important tool to improve early detection [13]. The objective of this study is to investigate the potential of *in vivo* Raman spectroscopy performed through small

diameter optical fibers to provide accurate diagnosis of atherosclerosis and the stability of atherosclerotic plaques. With the use of Raman microspectroscopy, the biochemical mappings of the transverse sections of coronary artery of different occlusions are reported, and different spectroscopic modeling methods are compared in this study.

The specific research objectives of this project were: i) to characterize the newly installed Raman microscope and to permit precise point-by-point mapping of pathological specimens, e.g. coronary arteries, ii) to design and carry out spectroscopic mapping of coronary artery specimens in order to correlate specific spectral features with histologically verified structural and compositional properties, iii) to analyze these data using established spectroscopic methods to obtain information about the chemical structure and composition of plaques, and iv) to utilize the results of this project to prescribe the development of optical fiber probe-based methods for future *in vivo* clinical use.

1.2 The Structure of an Arterial Wall and the Progression of Atherosclerosis

1.2.1 The Structure of an Arterial Wall

The artery wall is made up of three discrete layers (tunicas) that differ in structure [2,14]. The inner layer is a relatively thin layer called the intima, which is composed of a monolayer of endothelial cells joined to a basement membrane of proteins. The endothelium is coated with a layer called glycocalyx at the luminal surface. This layer is made up mainly of polysaccharides, glycosaminoglycans and glycoproteins, forming a smooth interface with blood. Under the view of a traditional light microscope, the intima is barely visible. The middle layer, called media, is thicker than the intima, and consists of almost entirely smooth muscle cells (SMCs) and variable elastin and collagen fibers.

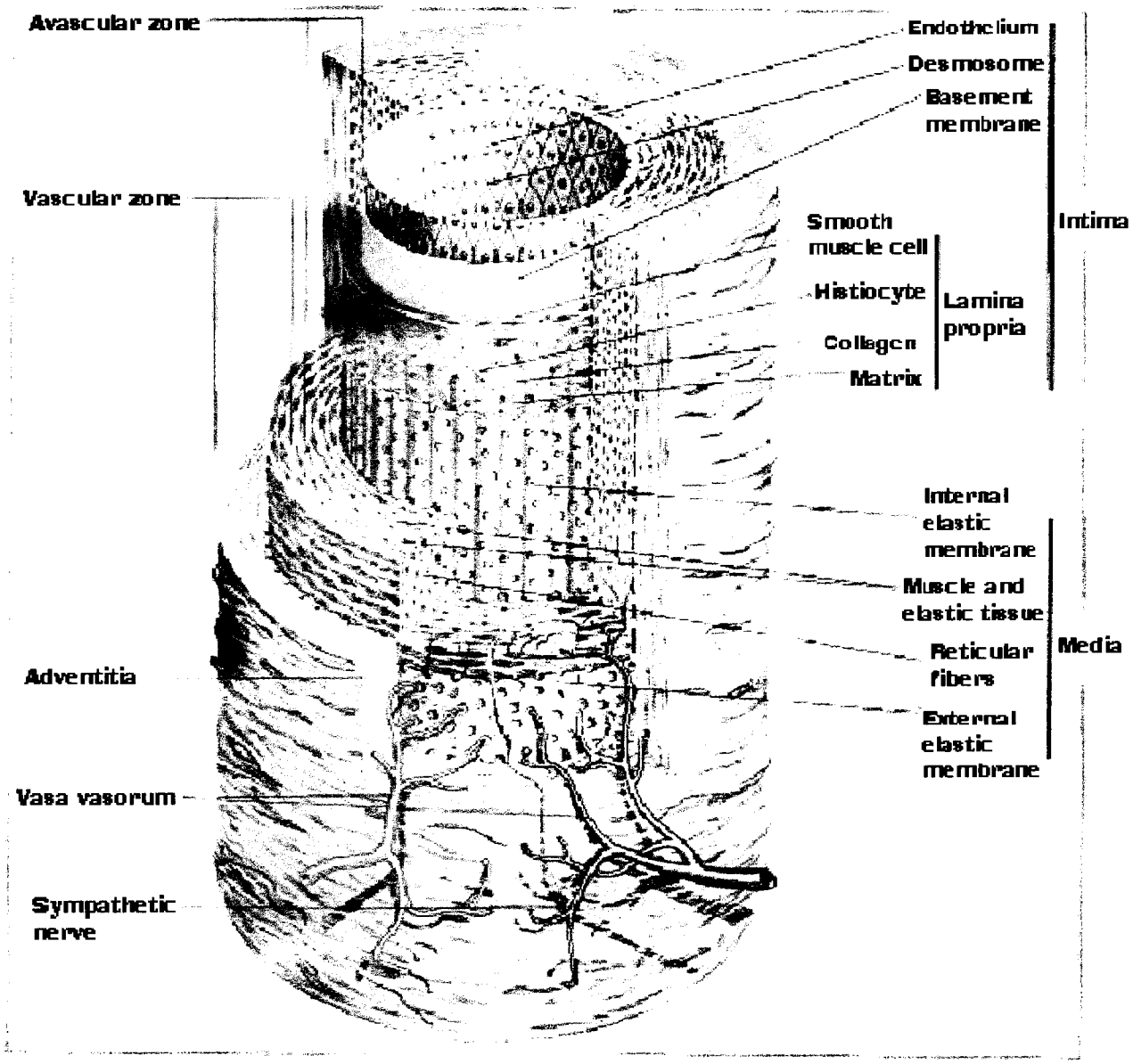


Figure 1.1 Schematic structural characteristics of the normal, disease-free arterial wall. Figure is reprinted from [15].

Separating from the intimal and medial layers is a thin layer of elastin called the internal elastic lamina (or internal elastic membrane) and is considered as a component of the media. The outer layer, called adventitia, is a relatively thin layer of connective tissue, composed primarily of loose collagenous fibers that run parallel to the long axis of the vessel. The media and the adventitia are separated by a boundary of elastic fibers called the external elastic lamina (membrane). The adventitia merges gradually with the surrounding connective tissue without a clear defined boundary. Atherosclerosis begins in and is predominantly a disease of the intima; secondary damage to the media may occur as the disease progresses [14].

1.2.2 The Progression of Atherosclerosis

The mechanism of artherogenesis is not completely clear and not yet fully understood, and there are a number of established hypotheses [1,2,15,16]. Many scientists believe the formation of atherosclerotic lesions is initiated by injury to the endothelium of the artery vessels. This initial step appears to be preceded and accompanied by inflammation, and it involves the breakdown in the integrity of the endothelial layer of the vessel, which permits infiltration of low-density lipoproteins containing cholesterol [1]. This leads to denudation of the endothelial layer, resulting in platelet aggregation at the intima surface and the release of platelet growth factors. These stimulate artery wall cells to produce other substances that result in further buildup of cells. Arterial smooth muscle cell and connective tissue proliferation as well as monocyte infiltration into the intima commences to repair the injury. Cells take up the cholesterol-rich lipoproteins to become foam cells. The progression of these lesions relies, in part, upon cellular uptake of lipid compounds, such as cholesterol and its derivatives [17], with a key intermediate step

being the precipitation of excess cholesterol in the form of crystals. As the lesion progresses, the cycle is repeated with old foam cells necrosing and localizing to the bottom of the lesion, near the internal elastic lamina, and becoming supersaturated with cholesterol. The buildup of necrotic material and mineralization in the form of calcium apatite crystallization leads to complicated advanced atheromas with hemodynamically important implications.

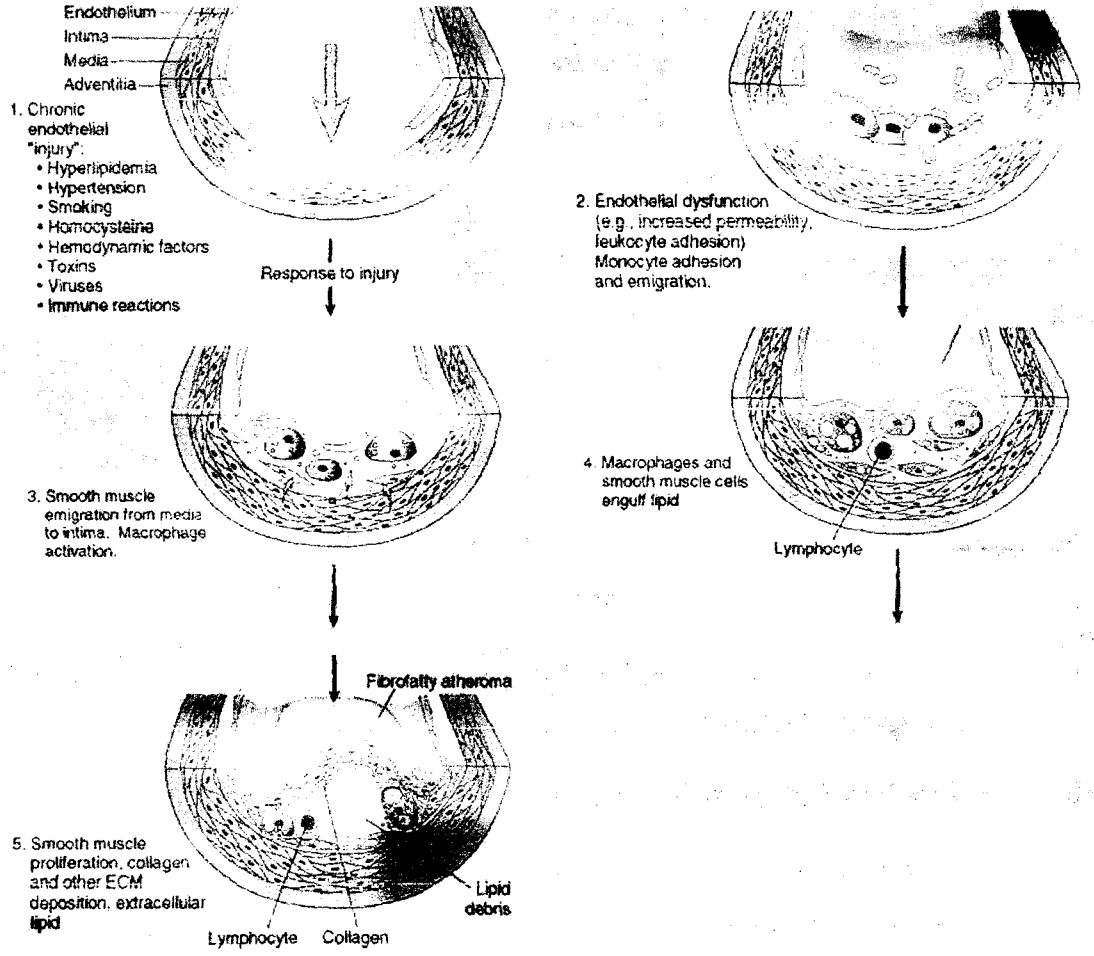


Figure 1.2 Schematic mechanism for intimal thickening. Figure is reprinted from [2].

Atherosclerosis is essentially an intimal disease, with changes beginning many years before they become clinically apparent [18]. Accumulations of lipid in the intima and associated changes begin in children and may lead to sympton-producing lesions starting at middle age. Atherosclerotic plaque's chemical composition and structure, rather than anatomy (narrowness of the blood vessel), determine whether an arterial lesion will progress, regress or rupture stability, and this view has been supported by many recent studies [19,20]. Nowadays, no reliable method has yet been able to certainly predict in advance disease progression and the risk of life-threatening complications such as thrombosis [10].

1.3 Current Artery Imaging Techniques

A variety of clinical imaging techniques have been used to characterize atheroscleroic plaques. Current diagnostic techniques are basically branched into two types, being invasive and noninvasive methods. Invasive techniques may sometimes restrict their use in patients.

1.3.1 Invasive Techniques

One of the most commonly applied invasive methods is the X-ray angiography, which employs an injected radioactive dye solution into the coronary arteries via a catheter (a thin, flexible, plastic tube), followed by taking x-ray images as the liquid flows through the arteries to determine the degree of obstruction [21]. Physician can identify blockages by tracing the flow of the radio-opaque solution. Even though angiography is in routine clinical use, it is not capable of determining plaque composition and hence its vulnerability.

IVUS (intravascular ultrasound), being another widely adopted invasive method, is based on transmitting and receiving high frequency acoustic waves delivered by the intravascular catheter to probe tissue intensity and images various layers of the artery wall and the structures within atherosclerotic plaques [22,23]. IVUS is able to provide information as to whether a lesion is atherosclerotic or due to a thrombus. Nonetheless, biochemical information is lacking, and thus determining the stability of plaques is again not feasible.

1.3.2 Non-invasive Techniques

Non-invasive imaging methods, like MRI (magnetic resonance imaging) and various sorts of tomography are the other two widely accepted diagnostic tools. MRI, based on the principles of nuclear magnetic resonance (NMR), has the potential ability to evaluate plaque composition. During examination, patients are subjected to a high local magnetic field, which aligns the protons in the body, with the use of powerful magnets. Images are then generated by the computer based on density of tissue [10,24]. NMR spectroscopy (^{13}C and ^{31}P) has been reported as capable of quantifying three components within the lesion: unesterified cholesterol, cholesterol esters and calcium phosphate salts [25,26]. Nonetheless, MRI is limited by availability, expense, spatial resolution and cardiac motion, plus the possibility of a complete biochemical composition profile is still under study.

Tomography can be carried out with a number of different methods, such as Cardiac CT (Computed Tomography), EBCT (Electron Beam Computed Tomography), and Cardiac PET (Positron Emission Tomography). The most intensively used method is EBCT. In EBCT, a beam of electrons, which is produced by x-ray radiation, passes

through the patient, and its attenuation is measured by a set detector rings [10]. Unlike MRI, EBCT imaging is relatively fast and thus eliminates artifacts caused by motion of organs yielding high spatial resolution [27]. In addition, EBCT can measure the amount of coronary calcification. The amount of calcification relates the amount of underlying coronary atherosclersis. However, the relation of calcification to the likelihood of plague rupture is still unclear [28,29,30].

1.4 Current Developing Diagnostic Techniques

Several new developing techniques, which do provide significant biochemical information of the plaques, have been focused on experimentally.

1.4.1 Fluorescence Spectroscopy

The earliest one has been fluorescence spectroscopy, by which, based on the amount of ceriods, arterial tissues were classified into three disease categories: non-atherosclerotic, non-calcified plaque, and calcified plaque [27,31]. Fluoresence spectroscopy involves the excitation of a molecule to its excited electronic state, followed by vibrational relaxation to the ground vibrational state of the excited electronic state, electronic relaxation to a possibly excited vibrational state of the ground electronic state, and finally possible relaxation to the ground vibrational state. Excellent detection limit and time response in fluorescence spectroscopy are achieved due to high quantum efficiencies for the fluorescence of many chromophores present in a molecule [27]. Nonetheless, since the fluorescence properties of many biochemicals are similar, fluorescence spectroscopy method lacks the ability to analyze tissue composition thoroughly.

1.4.2 Raman Spectroscopy

Raman spectroscopy is becoming the most appealing tool among others for the study and understanding of atherosclerotic disease progression [27,31]. One of the greatest advantages of Raman spectroscopy over other diagnostic techniques is its ability to identify the biochemical components of tissue and to determine their tissue concentrations in situ. As each Raman active biochemical exhibits unique (fingerprint) spectral features, and the relative contributions of these chemicals to the intensity of the overall Raman signals are proportional to their relative abundance in the tissue. Raman spectroscopy is a consistent analytical technique yielding objective and quantitative information about the lesion composition, leading to various medical applications [32]. A large number of biological molecules can be probed by using Raman spectroscopy, with many having characteristic fingerprint Raman spectra. Water exhibits weak Raman scattering while nonpolar bonds (organics) show strong Raman scattering, thus making Raman spectroscopy well suited for analysis of biological samples [13]. Recent technological advances in Raman devices should allow an *in situ* analysis and a better understanding of artherosclerotic disease. For example, Raman spectroscopy may offer as an attractive potential tool for surveying the biochemical changes that accompany the development of plaques, or as a method of monitoring plaque during laser angioplasty. (The use of lasers for the treatment of atherosclerosis by laser-initiated vaporization or ablation of the disease tissue site requires a prior knowledge of the composition of the site to be treated [10].)

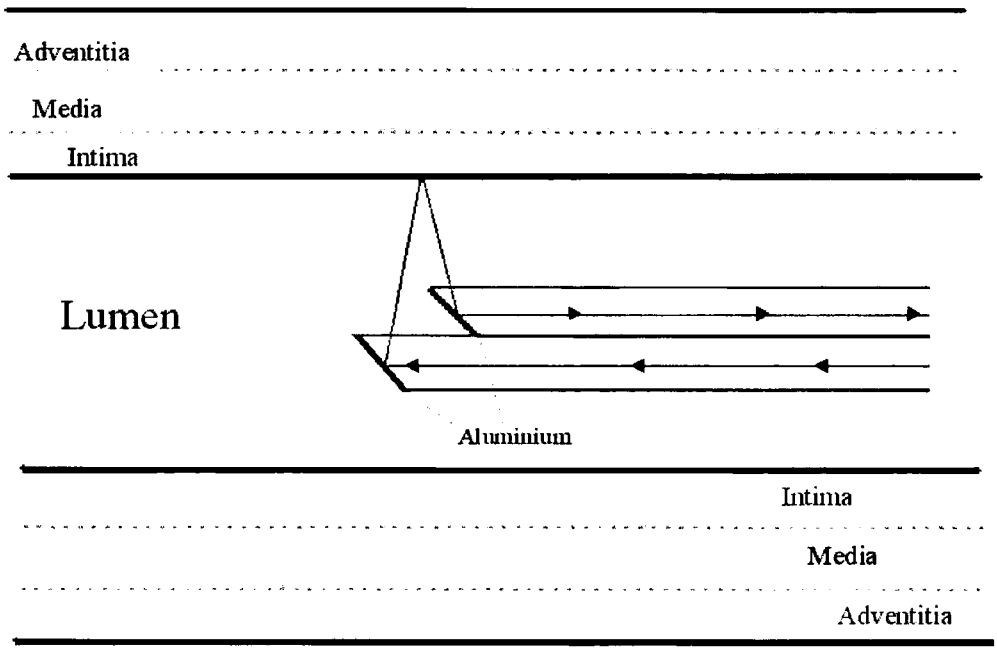


Figure 1.3 Schematic drawing of the proposed use of an *in vivo* optical fiber probe in the diagnosis of atherosclerosis. A right-angled excitation and collection probe geometry is necessary to examine the artery wall. Information about the arterial media and adventitia is not likely to be obtained unless the arterial wall is punctured. Hence, it is very important for Raman spectroscopy to be able to determine the disease progression by examining mainly the arterial intima.

Raman spectroscopy, like fluorescence spectroscopy, can be performed *in vivo* using remote optical fiber technology non-destructively, and requires the addition of no extrinsic labeling or staining, which may perturb the system under investigation [31]; the drawback in both cases is the absorbance of the light source by blood. However, Raman spectroscopy has several advantages over fluorescence detection. Raman spectra consist of sharp peaks specific for a given chemical structure, compared with the broad, generally featureless spectra of the fluorescence emission. In addition, the spectral

features are not expected to be significantly altered as the Raman radiation passes through any overlying tissue, since the bands span a very narrow wavelength range, about 2000cm^{-1} or less from the excitation wavelength. By contrast, fluorescence spectra, which span a wide range of wavelengths, can be altered by resorption as the fluorescent light passes through the tissue [10].

On the other hand, high background in Raman spectroscopy is usually a result of sample fluorescence, which is often intrinsic to the material of the sample and, therefore, unavoidably occurs as a result of using the laser to irradiate the sample [33]. Since fluorescence is often very much stronger than the Raman signal, it is necessary to minimize it in order to obtain good quality Raman signal. Raman spectroscopy has made remarkable progress in recent years, and the development in Raman instrumentation has brought resolutions to the interfering tissue fluorescence, and will be introduced in sections 1.6 and 1.7.

1.5 Principles of Raman Spectroscopy

The Raman effect was first discovered by, and subsequently named after the Indian physicist Sir Chandrasekhra Venkata Raman in 1928 [34]. When a molecule is illuminated with light, the scattering of light by the matter consists of mainly elastic scattering of the incident radiation (called Rayleigh scattering). The Raman effect refers to the inelastic scattering (about 1 part in 10^9) (Raman scattering) of the incident photons from a molecule. Raman scattering may result in increased (anti-Stokes) or decreased (Stokes) energy of the scattered light [35]. This low probability process of the incident electromagnetic field is associated with natural vibrations of molecular bonds, and that

the incident light is often referred as 'excitation' light because it excites the molecules into vibrational motion [34].

1.5.1 Classical View

Classically, this may be explained by the interaction of the periodically varying electric field of the incident radiation with the molecular vibration, mediated by the induced dipole of the molecule through its polarizability.

A beam of monochromatic light of frequency ν_{ex} is incident on a molecular bond in the material of interest. The time-dependence of the electric field intensity at the sample molecule is given by

$$E(t)= E_{ex} \cos (2\pi\nu_{ex}t) \qquad (1.1)$$

where E_{ex} is the amplitude of the wave in $J \cdot C^{-1}m^{-1}$.

When the oscillating field interacts with the polarizable electron clouds of the sample molecules, it induces a dipole moment μ_{ind}, which is proportional to the polarizability, α, of the molecule

$$\mu_{ind} = \alpha E = \alpha E_{ex} \cos (2\pi\nu_{ex}t) \qquad (1.2)$$

If the molecule is undergoing vibration, the polarizability may depend upon the internuclear distance, r. For small displacements, α can be expanded in a Taylor series about the equilibrium separation r_0, and ignoring second and higher order terms, α becomes

$$\alpha = \alpha_0 + (r - r_0) \ (\partial\alpha/\partial r \big|_{r=r_0}) + \ ..... \qquad (1.3)$$

where α_0 is the polarizability of the bond evaluated at r_0, the equilibrium position. The instantaneous internuclear separation r, varies harmonically about r_0 according to

$$(r - r_0) = r_m \cos (2\pi\nu_0 t) \qquad (1.4)$$

where r_m is the maximum internuclear separation and ν_0 is the natural frequency of vibration of the bond. Substituting (1.4) and (1.3) into (1.2), yields a first order (classical) description of the induced dipole moment,

$$\mu_{ind} = \alpha_0 E_{ex} \cos (2\pi\nu_{ex} t) + r_m E_{ex} (\partial\alpha/\partial r \big|_{r=r_0}) \cos (2\pi\nu_0 t) \cos (2\pi\nu_{ex} t) \qquad (1.5)$$

which, by considering the identity $A \cos\theta \cdot B \cos\phi = AB/2 \cdot [\cos (\theta - \phi) + \cos (\theta + \phi)]$, can be written as ,

$$\mu_{ind} = \alpha_0 E_{ex} \cos (2\pi\nu_{ex} t) +$$

$$r_m E_{ex}/2 \cdot (\partial\alpha /\partial r \big|_{r_0} \cos [2\pi (\nu_0 + \nu_{ex}) t] +$$

$$r_m E_{ex}/2 \cdot (\partial\alpha/ \partial r \big|_{r_0} \cos [2\pi (\nu_0 - \nu_{ex}) t] \qquad (1.6)$$

[35]

The classical treatment accurately predicts that there is Raman scattered radiation at higher and lower frequencies than the incident radiation accompanies the elastically scattered incident light at a frequency ν_0. It also demonstrates the requirement for a change in polarizability that causes Raman scatted light to be shifted both up (anti-Stokes

shifted) and down (Stokes shifted) by the frequency ν_{ex} of the molecular vibration. In addition, the classic view is capable of showing that the intensity of the Raman scattered light is directly proportional to the derivative of the polarizability with respect to vibrational displacement. In summary, Raman spectroscopy probes the vibrational modes of a molecule, and a vibration is Raman active if it changes the polarizability of the molecule.

1.5.2 Particle Theory View

In the particle theory view of Raman scattering, molecules are seen as being excited to non-quantized virtual energy states from where they immediately (10^{-15} seconds) relax to vibrational ground electronic state, with an associated change in the vibrational energy state of the molecule [34].

Generally, Raman scattering from a molecule may result in increased (anti-Stokes) or decreased (Stokes) energy of the scattered light, and is accompanied by a large mount of elastically scattered Rayleigh light. According to the particle theory view, Raman spectroscopy is an analytical technique based on the interaction of an incident monochromatic radiation with vibrational energy levels of molecules.

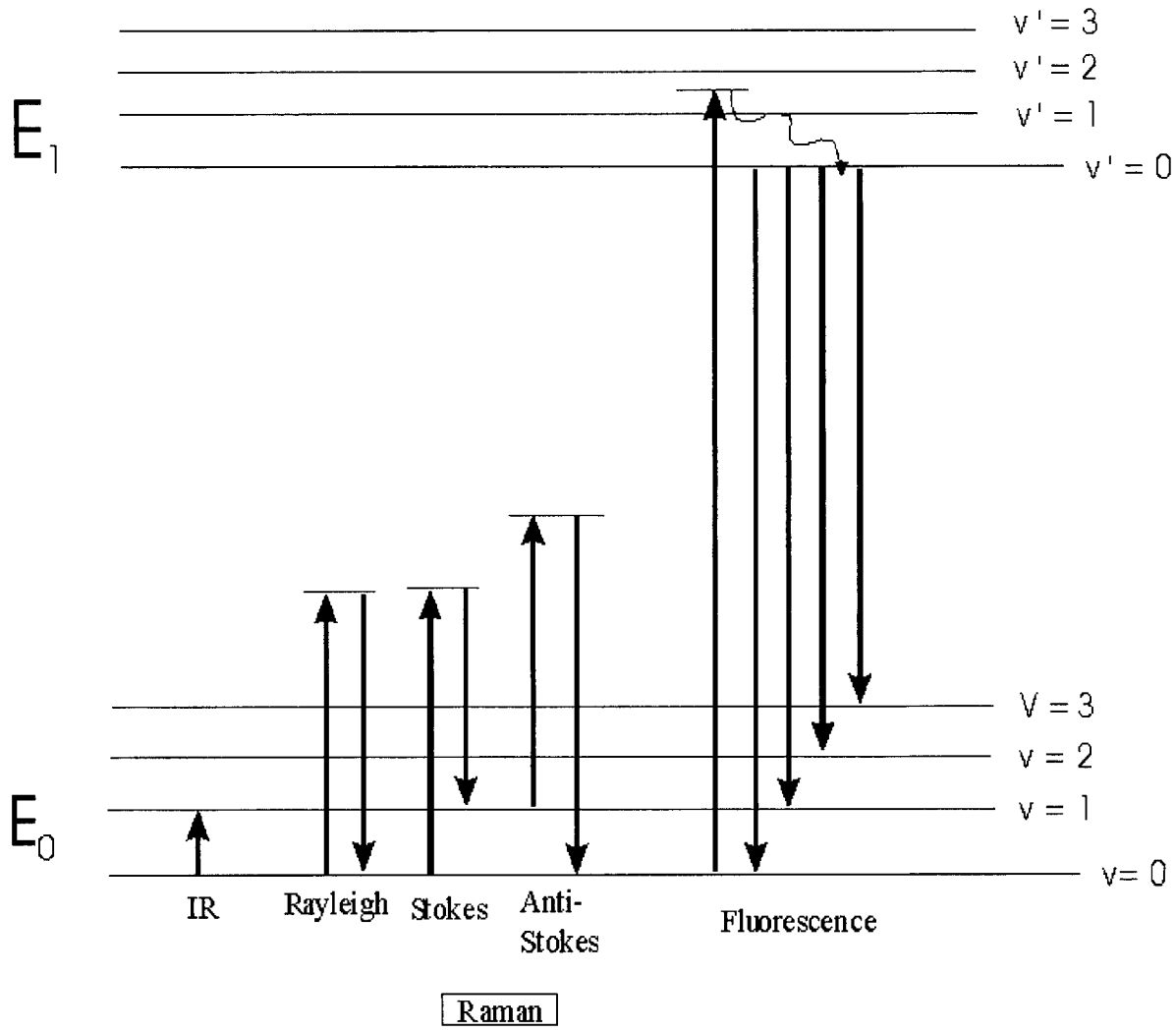


Figure 1.4. Energy level diagrams for Raleigh, Raman, and fluorescence events. E_0 and E_1 indicate ground and excited electronic states, v indicates the vibrational energy level of the ground electronic state, and v′ indicates the vibrational state of the excited electronic state.

1.5.3 Raman Spectra

Raman spectra are acquired by irradiating a sample with a powerful monochromatic laser radiation. The emitted radiation is of three types: Rayleigh, Stokes, and anti-Stokes scattering. Raman spectra are often plotted as signal intensity against the downward frequency shift from the excitation line (i.e. Stokes lines), and are conventionally recorded in wavenumbers (cm^{-1}).

Raman spectroscopy is sensitive to vibrational-energy-level differences [36]. Because each molecule has its own unique set of molecular vibrations, it results in species-characteristic inelastic scattering. Raman signals of each particular species show peaks, which are shifted by the characteristic vibrational frequencies of that molecule. Thus, Raman spectra provide chemical information at the molecular level.

1.6 Near Infrared (NIR) Excitation Raman Spectroscopy

Ultraviolet (UV), visible and infrared light (IR) can each be used for Raman excitation. However, if visible excitation is used in Raman spectroscopy, most materials exhibit strong, broadband fluorescence that can obscure the Raman spectra. Fluorescence decreases very rapidly at longer excitation wavelengths [37]. This is because the lowest excited electronic states of molecules, the state from which fluorescence originates, have energies corresponding to visible wavelengths, and so they cannot be excited by lower-energy near-IR light. Thus, any samples, including tissue, exhibit reduced fluorescence emission as the excitation wavelength increases into the near-IR region (>700nm). Therefore fluorescence interference in tissue Raman spectra can be greatly reduced by using near-IR excitation. In addition, near infrared (NIR) radiation does not harm the

organism, no electric potential application is needed, and the measurements are not influenced by electromagnetic fields [27]. Furthermore, NIR light can penetrate tissue deeply, of the order of millimeters, thus creating a 'diagnostic window' into tissue due to its relative minimal absorption [24].

Early NIR Raman bioanalytical spectroscopic studies almost exclusively employed Fourier-Transform (FT) spectrometers using a 1064nm excitation source, and demonstrated the promise of vibrational spectroscopy in bioanalytical applications [38,39]. The use of Fourier-Transform Raman spectroscopy with excitation at 1064 nm causes virtually no fluorescence [38]. Among those research, independent study has been focused on quantitifing the amount of a single or only a few constituents of the atherosclerotic plaques, such as cholesterol, calcium salts etc [27, 40-43]. FT Raman experiments require relatively high laser powers, which can damage or destroy light- or heat-sensitive samples, and/or long data collection time because of the low signal-to-noise (S/N) ratio of the commonly used InGaAs detectors and greatly reduced Raman cross-sections as compared with those at wavelengths in the 700 and 800 nm range [26].

The advent of low-noise, high quantum efficiency, cooled charge coupled device (CCD) detector technology and high-throughput dispersive monochromators makes lower excitation and collection time feasible. The advantage of using a CCD and dispersive monochromator was first reported in about around 1989 [27]. During the latest decade, many research groups have shown the promise of using Raman spectroscopy for early detection of numerous pathologies, including atherosclerosis, dysplasia, and cervical precancer, by the combined use of ~ 800nm NIR excitation, a dispersive grating and a CCD detector [12,33,36,44,45].

Several investigators have explored the potential of non-FT near-infrared (NIR) Raman spectroscopy for qualitative detection of atherosclerotic plaque *in vitro* [38]. Using a excitation wavelength below 1000nm range is not as good compared to FT-Raman using 1064nm in terms of fluorescence interference; nonetheless, Feld and colleagues have performed extensive laboratory work indicating that excitation wavelengths around 800 nm are optimal to minimize tissue fluorescence while best taking advantage of CCD detectors and short collection time. The first such application was done by the group in 1992, using an 810nm laser for *in vitro* studies of aorta tissues [27]. Subsequently, Feld and co-workers have conducted extensive *in vitro* studies on the quantitative biochemical analysis of coronary artery, biochemical-based diagnosis of artery, and very recently, morphological-based diagnosis of coronary artery using Raman microspectroscopy (RM) [19,20,45-47]. They have been able to show, using both biochemical-based diagnosis and morphological-based diagnosis, that *in vitro* Raman spectroscopy could classify the tissues into three groups, nonatherosclerotic artery, noncalcified plaque or calcified plaque, with a very low percentage misclassification [19,20].

In addition to applying non-FT NIR Raman spectroscopy by itself, non-FT NIR Raman combined with other techniques to assess the biochemical information have been investigated over the latest decade to study the progression of atherosclerotic plaque. For example, micro-PIXE (micro-particle induced X-ray emission), provides additional information on the elemental composition of the artery [48]. When combined with IVUS (intravascular ultrasound), Raman spectroscopy can be used to accurately quantify the chemical compositions of a particular tissue site of interest [42].

1.7 Confocal Raman Microscopy

Technological improvements in sensitivity and spatial resolution have drawn attention to Raman microspectroscopy (RM) as a useful tool in biomedical research. By means of Raman microspectroscopy, spectral information can be obtained from specific (small) regions in a biological sample. In addition, Raman microspectroscopy is a technique whereby spatially resolved Raman spectra can be obtained from an unprepared (unstained, unfixed and unlabelled) sample [49]. Together with its ability to record Raman spectra non-invasively and non-destructively, Raman microspectroscopy demonstrates to be a very unique identification technique for spatially resolved measurements in biological material, such as small volumes of coronary arteries.

High background in Raman spectroscopy is usually a result of sample fluorescence and background (the mounting material) fluorescence. Sample fluorescence is often intrinsic to the material of the sample and, therefore, unavoidably occurs as a result of using the laser to irradiate the sample. Since fluorescence is usually very much stronger than the Raman signal, it is necessary to minimize it in order to obtain the Raman signal. This can be achieved by using the confocal mode to suppression fluorescence from out-of-focus regions of the imaging area. Raman spectroscopic measurements were carried out using a confocal configuration in order to suppress signal contributions from the microenvironment of the selected volume, acquiring signals only from the part of the sample most strongly irradiated by the laser. Suppression of fluorescence from out-of-focus regions of the object under study in the confocal microscope typically operates with the use of a pinhole and a dichroic beamsplitter [50,51].

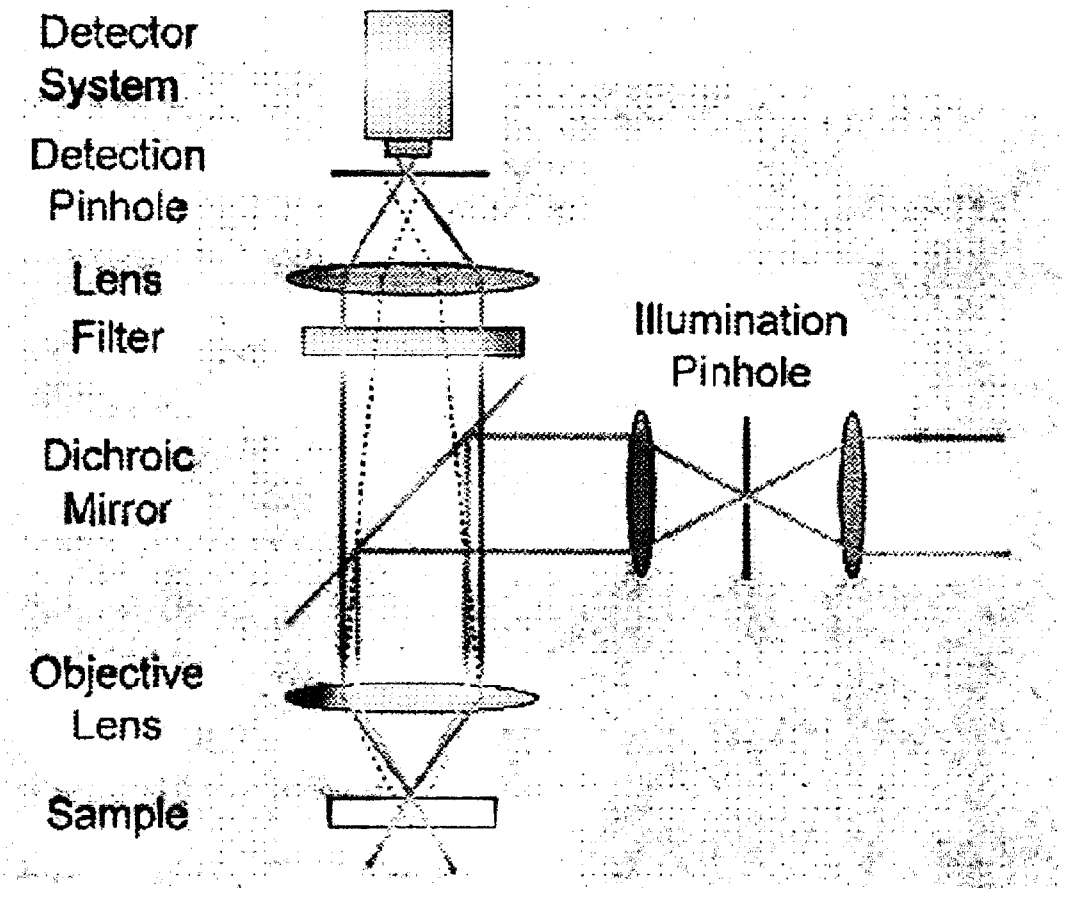


Figure 1.5 Schematic drawing of the laser light path in a typical confocal laser microscope, which operates with the use of a pinhole and a dichroic beamsplitter. Radiation not out of the focus (solid line) can pass the detection pinhole; radiation from other regions (dashed line) is blocked. Figure is reprinted from [51].

Thus, this 'optical sectioning' capability allows the reconstruction of objects in three dimensions, making RM be able to probe specific small volume [52]. Confocal spectroscopy is a method that enables spectra from well-defined regions of a sample to be collected with minimized spectral interference or contamination from the surrounding regions, compared to non-confocal operation. Together with the use of a precise,

automated stage, arterial mapping (by stepwise moving the sample under the laser spot while taking a spectrum after each movement) can be performed to obtain good-quality spectra from pre-selected area of interest on a transverse artery section.

The equipment set-up, system characterization, experiment design, development and performance of a confocal Raman microspectrometer (CRM) that enabled our group to record high quality non-resonant Raman spectra from coronary artery sections are reported here.

Chapter 2

Instrumentation and Confocal Mode Setup

2.1 Instrument Setup

A single frequency green 532nm output Verdi^{TM} V-10 solid-state diode-pumped, frequency-doubled Neodymium Vanadate (Nd:YVO_4) laser (Coherent Laser Group, CA), which has a maximum output power of 10 Watts, was used as a pumping source. Near-infrared CW (continuous wave) excitation light was generated by a titanium:sapphire ($TiAl_2O_3$) Mira Model 900 laser (Coherent Laser Group, CA), which is tunable from approximately from 700 to 1000nm. The wavelength of the laser output was adjusted to 785nm. The excitation wavelength was measured with a Miniature Fiber Optic spectrometer model SD2000 (Ocean Optics, Inc.). The laser output was band pass filtered with a holographic laser bandpass filter model HLBF-785.0 (Kaiser Optical Systems, Inc., MI), which has a throughput transmission >90% and allows a narrow bandwidth of less than 2nm.

All studies were performed using a Renishaw Raman Imaging Microscope System 1000B. The spectrometer box consists of a 1800 lines/nm grating, which yields a spectral resolution of approximately 1 cm^{-1}/pixel [50]. Instead of employing a traditional beamsplitter, the Renishaw Raman spectrometer optimizes coupling efficiency with the use of holographic notch filter (HNF). By eliminating the use of a 50/50 or 70/30 beam splitter, the holographic notch filter delivers >90% of the Raman light to the grating spectrometer [53]. The advantage of this high efficiency coupling is that relatively lower power is required from the laser and that relatively shorter spectral measurement time is needed. In addition, the holographic notch filter rejects Rayleigh light from the sample

and removes the complication of stray light from the laser line for measurements close to the excitation wavelength. The HNF also allows the possibility of simultaneous Stokes and anti-Stokes Raman signal collection.

The spectrometer box is coupled to an infinity corrected Leica DMLB Microscope fitted with phase contrast system. The microscope has 5x, 20x and 50x objectives. The microscope stage equipped is an automated motorized xyz stage with focus motor (ProScan Series, Prior Scientific Inc., MA). x and y axes resolution (i.e. step size) can be as small as 0.1μm, while z axis resolution is 0.01μm. The sample can be directly viewed through the microscope binocular. Simultaneously, the images of the samples can be displayed on the screen of the computer as live colour video with the microscope Video Viewer and be saved.

Unstained sections or solid form of chemicals were placed on the microscope stage for spectroscopic experiments performed at room temperature. Initial selection of the area was performed with phase contrast microscopy at 5x magnification. The Raman spectra were collected under the 50x magnification. The excitation power from the laser line was measured with a laser power meter (Gentec TPM-300CE) and recorded before Raman signal collections. The details of Raman spectra acquisition will be described in Chapter 4.

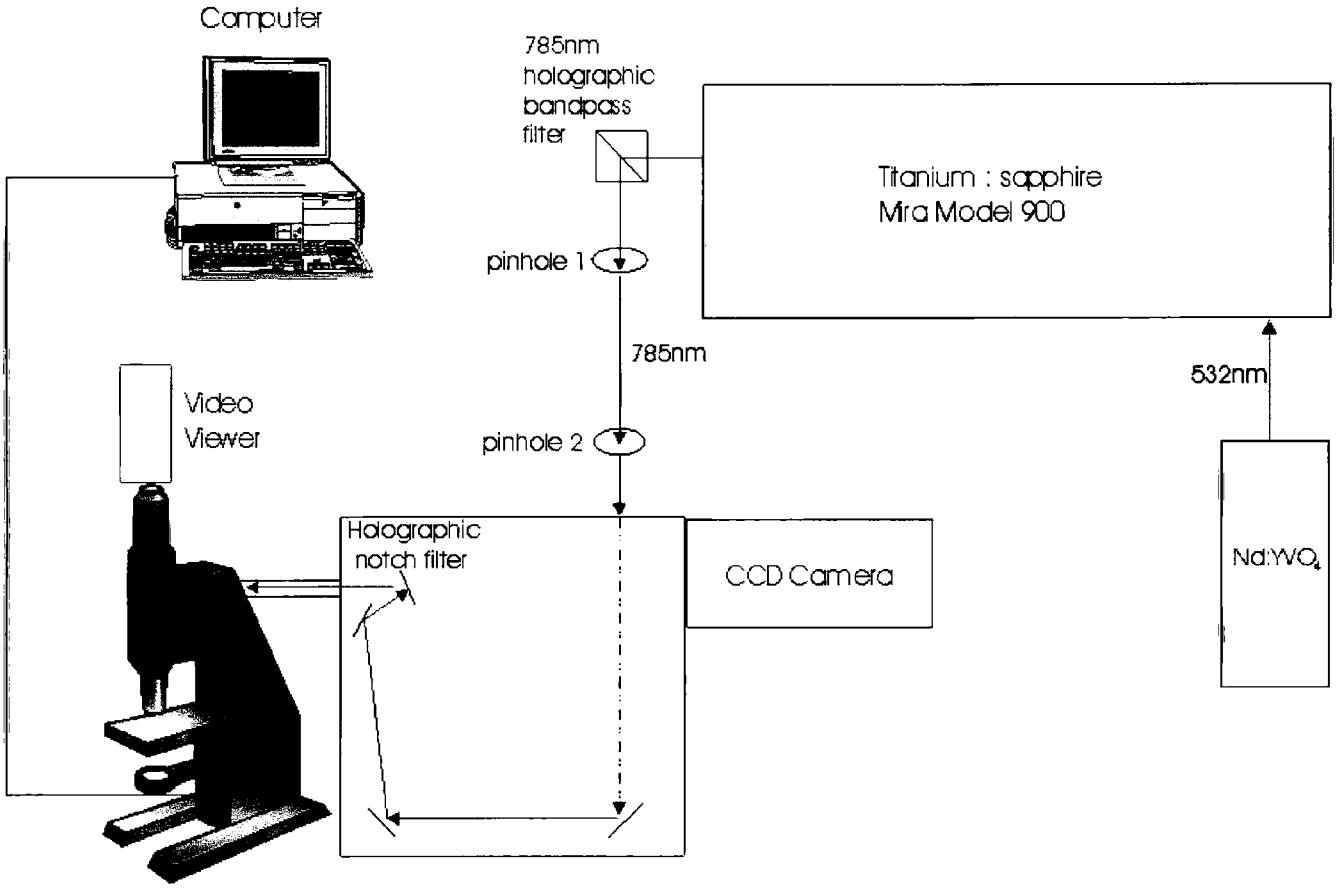


Figure 2.1. A schematic representation of the lasers and the Renishaw Raman Imaging Microscope system setup. An 785nm laser light is delivered by a Ti/sapphire laser system. Before the laser light was directed to the microscope unit, two pinholes were set-up and in fix position to ensure the laser beam path was exactly unchanging each time the laser was operated. Three mirrors were located in the laser path between the bandpass filter and the first pinhole iris. Fine adjustments of the mirrors were required to ensure a consistent beam path. Only the excitation beam path was indicated above. The arrangement of optics and holographic notch filter within the microscope unit, and the scattered light collected from the sample are further described in figure 2.2(a).

Signal detection is achieved through the use of a sensitive charge coupled device (CCD) array detector camera (thermoelectric cooling to –70°C), with instrument control performed by a personal computer. The software package consists of two programs working together, instrument control software (Renishaw WiRE™ - Window® based Raman Environment, software version 1.3.15) and data analysis software (GRAMS/32® - Graphic Relational Array Management System, version 4.14 from Galactic Industries Corporation). The optical alignment of the instrument was optimized using the 520cm^{-1} Raman signal from a piece of silicon semiconductor. Calibration of the wavenumber offset was done every time before the signal collection using marked positions on a piece of silicon semiconductor and on a Tylenol™ tablet.

2.2 Renishaw Confocal Mode Setup

All Raman spectroscopic measurements were carried out using the confocal configuration. Setting the system in this way ensures that light from the region of the sample which is in best focus to the Raman system is accepted, while light from adjacent regions, which is be out-of-focus, is rejected. In this way, setting the system to be confocal "concentrates" the view of the spectrograph on the region of the sample that is of greatest interest, i.e. optimizing spectrometer performance.

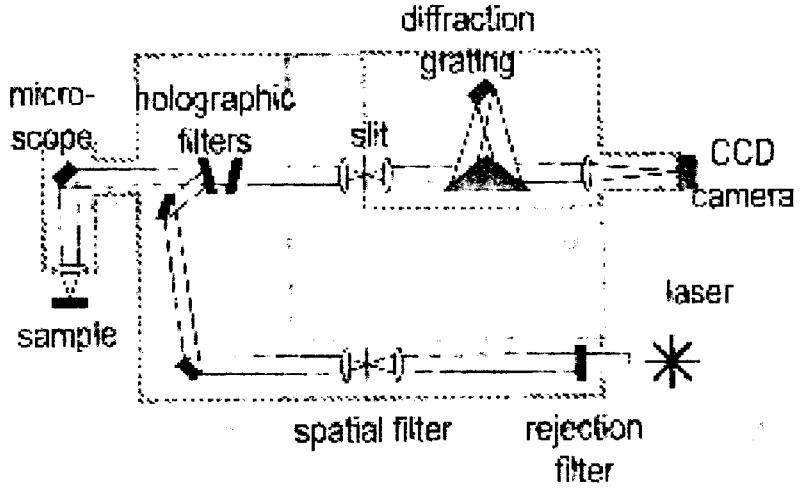
However, in contrast to more conventional pinhole arrangements as described at the end of Chapter 1, the Renishaw Raman Microscope system employs the combination of a variable slit width and a pre-selected number of pixels from the CCD to act as the primary aperture in the typical confocal microscope system. The slit acts as a one-dimensional spatial filter for the light entering the grating spectrograph. The mode of data

collection, which depends critically on changing the area of the CCD used to capture data (the capture box), provides an additional spatial filter at the 90 degree angle under confocal measurements as seen in figure 2.2.

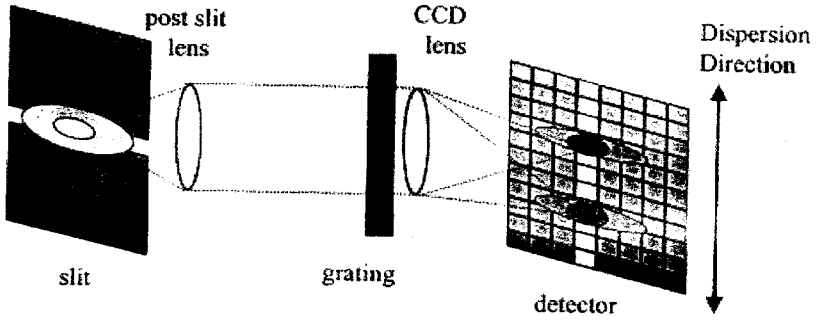
The confocal mode of operation was achieved by i) decreasing the slit width to 5μm, and ii) reducing the area of the CCD image used for capturing grating spectra to 4 x 576 pixels. At the confocal setting, the 50x magnification lens resulted in 5μm sampling depth (or penetration) [53,54].

Confocal Raman spectroscopy is a method available to obtain spectra with increased lateral resolution and well-defined, decreased depth of field, by restricting the field of view of the system. The lateral diameter of the "confocal volume" (or the sampling volume) is approximately 2 microns and the depth is approximately 5 microns, determined by Renishaw [53,54].

a)



b)



c)

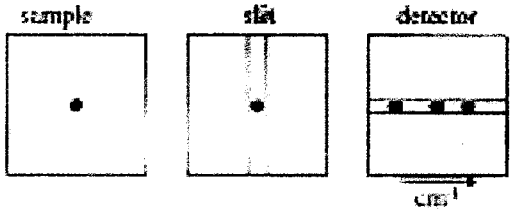


Figure 2.2 Optical configuration under the confocal operation mode of the Renishaw 1000B system. (a) confocal optics, including the holographic notch filter, and the use of a slit instead of a traditional pinhole. (b) Silt width controls the amount of signal being collected in x-direction, while CCD pixel binning determines the amount of pixels to be registered in the y-direction. (c) The combined use of the slit and CCD pixel binning act as a typical pinhole to reject signal from out-of-focus contribution. Figures (a) and (b) were reprinted from [53], and (c) was reprinted from [54].

Chapter 3

Coronary Artery Specimens

3.1 Sample Preparation

Sectioned and mounted coronary artery specimens were obtained through collaboration with practising physician and researchers specializing in cardiovascular diseases at St. Paul's Hospital, Department of Pathology, Vancouver, British Columbia.

The arteries are collected from autopsy cases, approximately 2 to 5 days after death, and the samples are collected via informed consent through the next-of-kin. The samples are excised from the heart in a cross-sectional manner and embedded in OCT compound using plastic molds. OCT compound is a company trade name and stands for "Optimal Cutting Temperature", and is an embedding media used for frozen tissue. It is made by Tissue-Tek and it contains: 10.24% w/w polyvinyl alcohol, 4.26% w/w polyethylene glycol, and 85.50% w/w nonreactive ingredients. They are then frozen on a dry ice / isopentane mixture to freeze at -80°C. The amount of luminal narrowing (atherosclerosis) is estimated and documented. The samples are stored in an -80°C freezer.

Transverse tissue sections 20 and 25 μm thick for light microscopy and Raman microspectroscopy were sectioned using a Reichert-Jung cryostat (Leica, model Cryocut 1800). There is a micrometer screw on the machine that is used to set the section thickness. The cryostat has been serviced regularly, and it is assumed that the micrometer is also serviced to provide accurate thickness.

Serial stained and unstained arteries were placed on regular glass slides wrapped with the non-shiny side of commercially available aluminium foil. Problems were

initially encountered with the thin coronary artery sections as the glass slide material was found to produce a significant underlying fluorescence background which degraded the signal to noise ration (SNR). The problem was mitigated by using a metal surface on which the tissue mounted instead. The stained specimens are for light microscopy while the unstained specimens were for mapping experiments, in which the actual Raman signals were collected. The stain used to visualize the artery is called "Hematoxylin and Eosin" stain which is a routine stain used in histology. The hematoxylin is the blue stain and the eosin is red. All sectioned artery specimens were stored at -80°C until use.

3.2 Coronary Artery Specimens Investigated in this Study

The amount of artery cross-sectional area narrowing was roughly determined by visual grading of the surface area of an artery involved in disease, despite the limitations such as grader bias and consistency.

Four differently occluded arteries were provided for the present study, they were 0%, 0-25%, 50-75% and 100%. The atherosclerotic plaques investigated in the study were all stable. Serial unstained sections of 20 and 25 μm thick were mapped to collect Raman spectra. The features on the stained sections should correspond well to those on the serial unstained sections, except the shape of the artery may have been distorted slightly during sectioning and when being placed on the slides.

Photos of the stained coronary artery specimens were taken using the microscope video viewer under 5x magnification.

Figure 3.1 Photo of a cross section of the stained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 50-75%. The spasm shows eccentric atherosclerotic plaque with a small tongue-like intimal proferation.

Figure 3.2 Photo of a cross section of the stained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 100%. The spasm shows concentric atherosclerotic plaque with the intimal proliferation almost totally occludes the lumen.

Figure 3.3 Photo of a cross section of the stained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 0%, showing normal intima.

Figure 3.4 Photo of a cross section of the stained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 0-25%. An early lesion, with intimal proliferation apparent on the top and left side of the artery.

Chapter 4

Raman Signal Acquisition

4.1 Raman Spectra of Biochemicals

Reagent grade commercial biochemicals were used to obtain the Raman spectra for use as the basis spectra. A model, which will be discussed in Chapter 7, was developed to determine the relative fit contribution of these biochemicals to the Raman spectra of transverse sectioned artery specimens. A total of eighteen biochemical components were selected, including cholesterol (Sigma); cholesterol esters: cholesteryl oleate (Sigma), cholesteryl linoleate (Sigma), cholesteryl palmitate (Sigma), cholesteryl stearate (Sigma), and 22(S)-hydroxycholesterol (Sigma); triglycerides: triolein (Sigma); phospholipids: L-α-phosphatidylcholine (Sigma); calcium salts: hydroxyapatite (Fluka) and calcium carbonate (Fisher); β–carotene (Sigma); proteins: collagen type III (Sigma), actin (Sigma), myosin (Sigma), tropomyosin (Sigma), elastin (Sigma); and sugars: d-glucose (Gibco) and sucrose (Fisher). Raman spectra were collected for these eighteen biochemicals as they represent the most common biochemical species found in normal arterial tissue and in atherosclerotic plaques [39,47].

The spectral range selected was from 300-1800 cm^{-1} wavenumber shift. Seventeen out of eighteen biochemicals were available as solid state, and the spectra were collected microscopically at random locations on the surface of each substance, with no cover slip, and in complete darkness (room light and light from the computer monitor were all shut down). However, spectral collection from different locations were repeated

and compared in order to avoid any variation due to the non-homogeneous nature of individual crystal or due to impurities.

For the liquid, a drop of sample was placed on the aluminium-wrapped slide surface. Myosin, which is commercially available in liquid form, was focused under the microscope (50x objective) such that any bubbles in the liquid were out of the field of view. A power of ~ 40 to 50 mW on the substance was used. An integration time of 30 seconds with no binning of data (i.e. spectrum resolution remained ~ 1 pixel/cm^{-1}), no cover slip, and in complete darkness were the chosen parameters.

The collection of Raman spectra for β-carotene, which is a very strong Raman scatterer, required a laser power of less than 10mW at the sample, no binning of the data, and an integration time of the minimum required 30 seconds to prevent saturation of the CCD array occurred.

On the other hand, the solid proteins (collagen, actin, tropomyosin, and elastin) are very weak scatters, and the Raman signals captured were weak of a short integration time. Thus, in order to bring out the signals from the background noise when collecting spectra from these four proteins, a laser power of ~ 40 to 50 mW on the sample was used, with a five-point vertical binning of the data (i.e. spectrum resolution became ~ 5 pixel/cm^{-1}), at an integration time of 100 seconds, and the accumulation of 5 spectra to produce final spectra with observable peaks.

To collect Raman spectra for the rest of the biochemicals (i.e. cholesterol, cholesterol esters, triglycerides, phospholipids, calcium salts, and sugars) a power of ~ 40 to 50 mW on the substance, and an integration time of 30 seconds with no binning of data was sufficient.

4.2 Raman Mapping of Coronary Artery Specimens

Unstained artery sections were placed on the microscope stage for spectroscopic experiments performed at room temperature and in complete darkness. No cover slip was used for these spectroscopic measurements. Initial selection of the area was performed with phase contrast microscopy at 5x magnification. The Raman spectra were collected under the 50x magnification. The laser output was focused onto the sample using the focusing control of the microscope.

The artery specimens were of two different thickness 20μm and 25μm. Since the sections were thin, they looked almost transparent when viewed under the microscope. However, because they were placed on the surface of the aluminium foil, we were able to focus initially on the aluminum surface and then move away from the aluminium surface a known distance to make sure the collection volume was of the artery that contains minimal contribution from the aluminium surface.

The best location to collect signals from the artery sections was determined to be a location about half the thickness of the specimen above the surface of the aluminium surface. For the 20 μm thick sample, the stage was moved an extra distance of 10 μm away from the objective lens after initial focusing on the aluminium surface and before signal collection started. For the 25μm thick sample, a defocus of the aluminium surface was required as well, and an extra 12μm distance between the objective and the stage was needed. In both cases, the artery signals collected were showing a reasonable amount of reflection from the aluminium foil and yet the signals from the biochemical constituents were desirable.

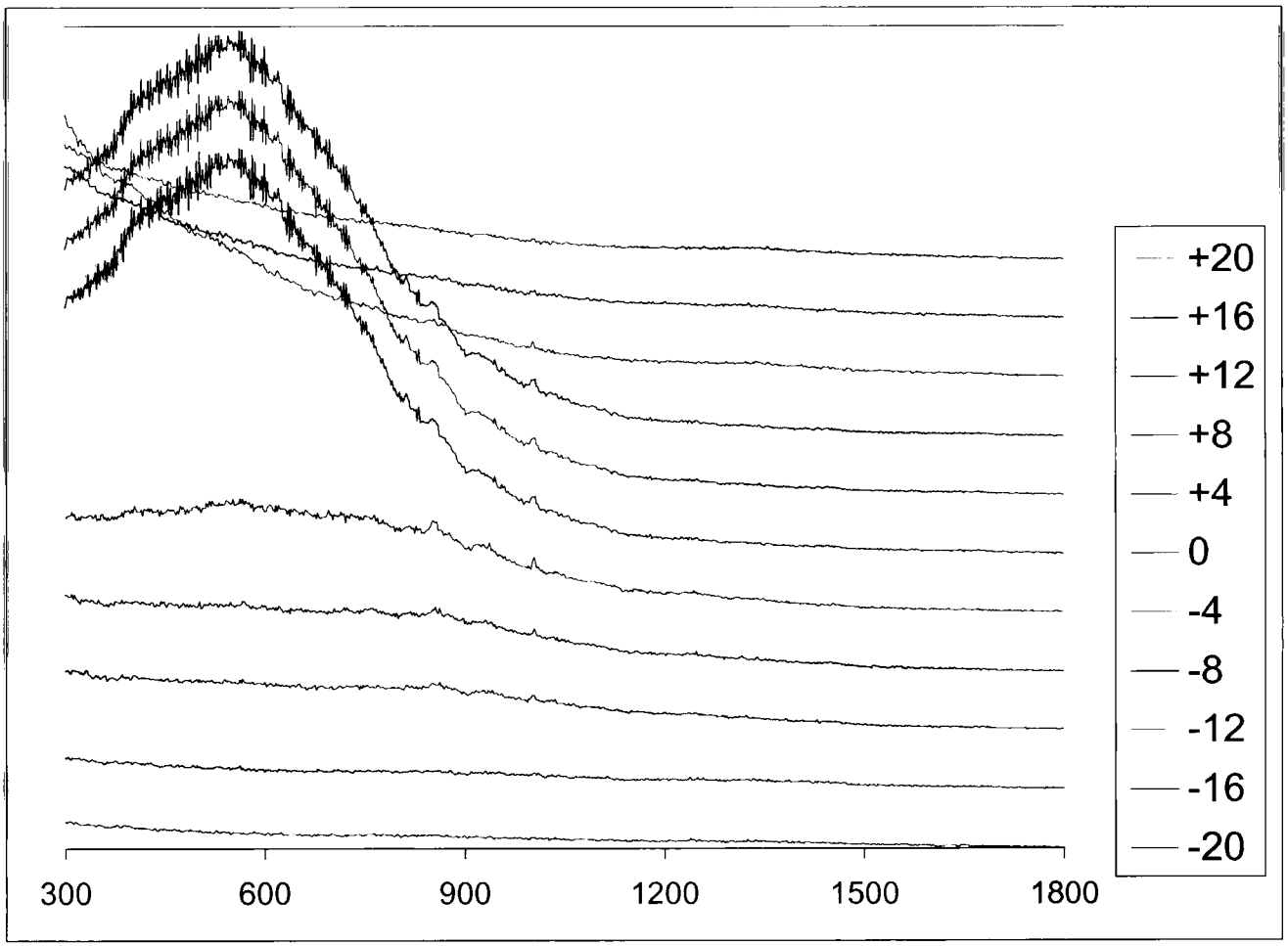


Figure 4.1 A series of artery spectra obtained with equivalent parameters. They indicate the change in S/N with the effect of varying distances from the objective's focal point, which is the aluminium surface. (+) represents moving the stage closer to the objective, while (-) represents moving the stage away from the objective. Numbers indicated in the legend represent distance in micrometer (μm).

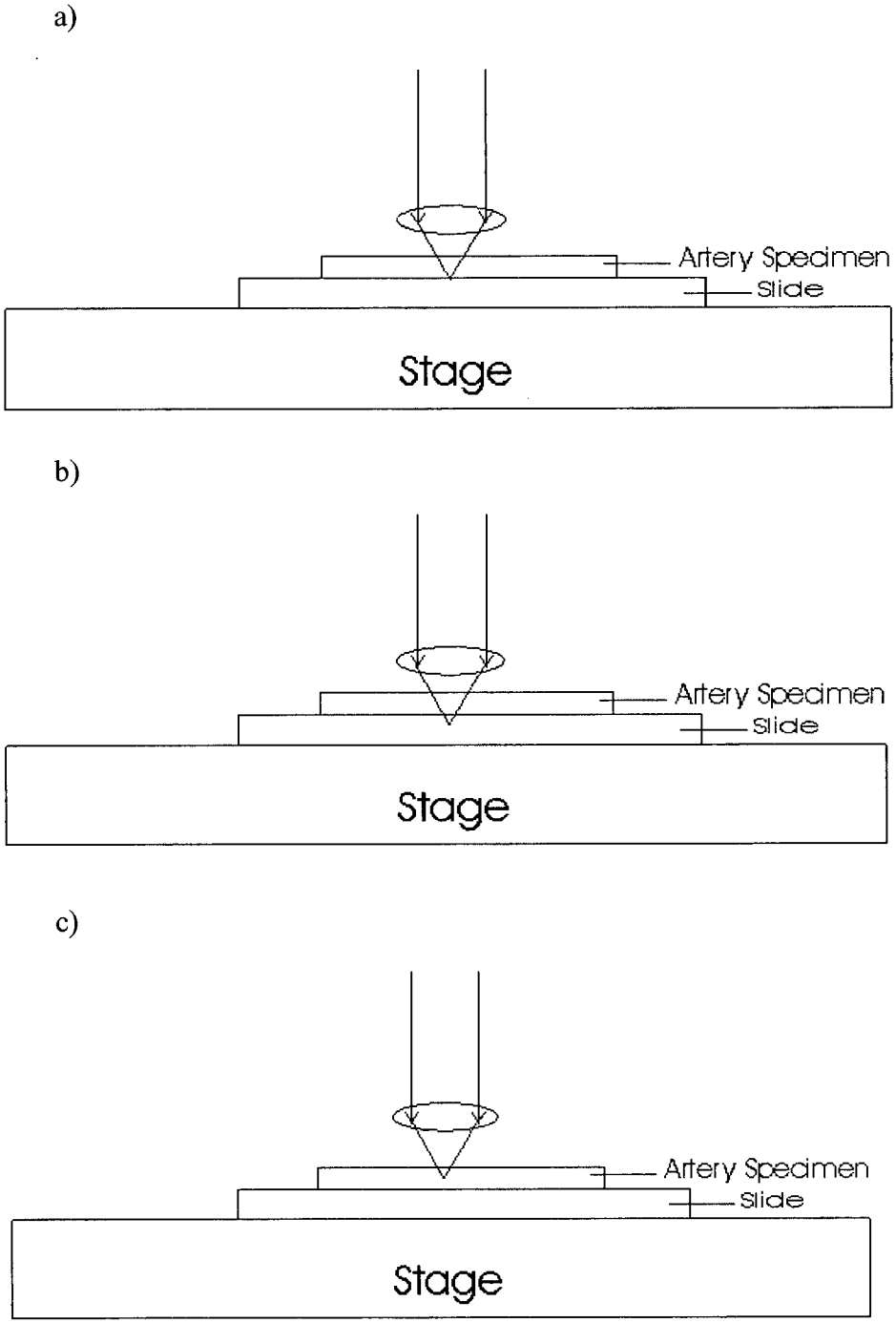


Figure 4.2 Schematic representation of the various locations of the focal point of the objective when the stage was (a) remained unmoved (focus was on the aluminium surface), (b) being moved towards, and (c) being moved away with respect to the objective. The condition described in (c) was employed for the Raman artery spectrum collecting.

Mapping of a specimen means the stepwise moving of the sample under the laser spot while taking a spectrum after each movement. All the artery samples in this study were mapped for selected areas of interest. Raman spectra were collected in 20μm to 100μm step sizes over those areas, covering the three layers of the coronary artery (intima, media and adventitia). The sample was irradiated with a power of ~ 80 to 100 mW on the sample, together with an integration time of 30 seconds with no binning of data (spectral resolution ~ 1 pixel/cm^{-1}). The exact areas where the different artery sections were mapped will be shown in the Chapter 6 and 7, when correlation and composition from different parts of the artery are mentioned.

Before the start of each mapping experiment, the arterial section was placed evenly flat on the stage. By manual adjustment of the height of the four corners of the stage, the focus on the aluminium surface was necessarily kept constant over the selected area of interested. The specimen was mapped automatically, without the loss of laser focus, hence without having to manually focus the laser after each stage movement. After the mapping was finished, the spectra having "strange" or unexpected signals were recollected. Random locations were also selected and signal collection was repeated to determine whether biochemical changes on the sample due to enzymatic degradation happened within the mapping period when the specimen was left in room temperature. It was shown that no significant changes were seen in Raman spectra within the mapping experiments.

Chapter 5

Post Acquisition Processing of Raman Spectra

5.1 Post Acquisition Data Processing

5.1.1 The Need for Raman Spectra Baseline Removal

Raman signals are generally weak and often superimposed on a large fluorescence and Rayleigh background/baseline. More significant and frequently performed manipulations are signal-to-noise ratio (SNR) enhancement and deconvolution of the instrumental line shape [55,56]. SNR enhancement in Raman spectroscopy is important because of the inherently weak nature of the Raman effect and the frequent existence of a large fluorescence background, which degrades SNR due to shot noise. In addition, the distortions caused by noise often interfere with the interpretation and quantification of Raman data: weak peaks may be obscured, and the measurement and calibration of peak volumes can be problematic. The error is even more pronounced when the baseline is strongly nonlinear and has a large slope due to the Rayleigh scattered light.

The study of biological composition by means of Raman spectroscopy relies on the interpretation of the spectrally overlapped multiple spectra. The unprocessed spectra are subjected to baseline distortions caused by fluorescence, etc, and thus baseline correction is essential to obtain more meaningful, interpretable results. Flattened baselines are needed to measure peak volumes and to permit the detection of the weak Raman signals. Spectra with true peak magnitude and shape are obtained after the removal of the distorted baseline [57]. This raises the issue of spectral processing, and the need for good baseline subtraction. The accuracy of determining the true peak magnitude and shape depends on the accuracy of the baseline correction. Baseline removal permits

the separation of the two components, spectrum and baseline, and the extraction of enhanced Raman spectra.

5.1.2 Automated Baseline Correction

There exist numerous techniques for baseline removal from spectra, because the problem is ubiquitous within spectroscopy [58]. Methods are scattered across many fields of research and there exists no recent review article [58]. Nevertheless, Vickers and Mann [59] summarized some baseline treatment methods being employed to Raman spectra.

In many cases, baselines are best removed simply by manually selecting for the baseline [58,60]. The manual correction usually shows good results, but it is too time-consuming for the whole set of spectra, for instance in mapping of the arterial wall. Therefore a reliable, automatic method is needed for a large number of spectra, otherwise, the processing time of 2D spectra may easily exceed the acquisition time. It is also beneficial to use numerical tools that minimize the need for judgement calls and permitting reproduction of the results by others. The ability to extract quantitative peak intensity and line shape information from spectra can in many cases be limited by the ability to remove an underlying continuous baseline, and in such cases automated numerical techniques of baseline estimation are clearly necessary.

For many cases, the sources of artifacts are unknown and the shapes of the distortions are unpredictable [61]. The classic approach to remove the artifacts is to fit a predefined functional form to a subset of points, which lie on or near the baseline [62]. Estimates of the baseline values for the remaining points are then obtained by interpolation and extrapolation. The approximating function serves as a model of the distorted baseline and, when subtracted from the spectrum, yields a flattened baseline.

The effectiveness and accuracy of automatic baseline correction is confined to the design of robust peak-picking algorithms [62]. Numerous approaches have been developed to correct baseline defects during post-acquisition processing. However, the correction of baseline offsets becomes much more difficult in the presence of baseline artifacts with arbitrary magnitudes and shapes, given the wide range of possible baseline distortions and peak shapes. For this reason, two baseline removal methods were employed and compared in this study; the first method estimated the baseline with the use of natural cubic spline function, while the second algorithm was based on a method employed by Dietrich *et al.* in NMR spectra baseline removal [60]. Each set of the baseline-subtracted spectra was used for 2D correlation analysis (Chapter 6) and chemical composition analysis (Chapter 7).

Baseline correction of all Raman spectra collected in this study was performed with the two customized automated algorithms using the software HiQ version 4.1 (National Instruments Corp., Austin, TX). The programming codes are included in Appendix A and Appendix B. Cosmic ray removal, subtraction of the broad background originating from the OCT or from the optical elements in the laser delivery pathway, and removal of the tissue fluorescence/Rayleigh, are included by both methods, which will be discussed in the sections that follow. The two algorithms are fully automated and make no assumptions about the shape or functional form of the background. The positions of the Raman signals and their height in arbitrary units were determined with respect to the baseline by the program.

5.2 The Use of the Natural Cubic Spline Functions

5.2.1 Theory of Cubic Spline

An outline of a mathematical approach in which cubic spline functions are used to remove baseline artifacts in Raman spectra is discussed in this section.

The description of the baseline by cubic splines is more accurate as compared to the straight line. Spline interpolation uses several lower degree (3^{rd} degree in cubic spline) piecewise-continuous polynomials for interpolation of a data set, avoiding problems associated with higher degree single polynomial interpolation. Spline interpolation requires the interpolant to pass through the data set used to compute the interpolation. In this case, nodes are selected outside the peak regions in the experimental curve, and the function is interpolated by cubic splines between the nodes selected [63-65].

First, let $y_i(x_i)$ be the given ordinates at x_i ($i = 1, \ldots, n$). A spline approximation of the baseline requires the use of a cubic spline function f(x),

$$f_i(x) = a_i + b_i(x-x_i) + c_i(x-x_i)^2 + d_i(x-x_i)^3 \qquad \text{for } x_i \le x \le x_{i+1} \qquad (5.1)$$

$$(i = 1, \ldots, n-2)$$

where a_i, b_i, c_i and d_i are called spline coefficients, which can be solved from the sewing conditions for the neighboring splines at the node points, using the conditions described below.

The Fitting of the spline function requires a few conditions to be satisfied:

1) spline value f(x) should be equal to the initial function values f(x) at the node points (called Lagrange condition),

i.e. $f_i(x_i) = y_i \qquad i = 1, \ldots, n\text{-}1$

$f_{i-1}(x_i) = y_i \qquad i = 2, \ldots, n \qquad$ etc, (5.2)

2) The first and second derivatives of splines should be continuous at the node points,

i.e. $f'_{i-1}(x_i) = f'_i(x_i)\ , \qquad i = 2, \ldots\ldots, n\text{-}1$

$f''_{i-1}(x_i) = f''_i(x_i)\ , \qquad i = 2, \ldots\ldots, n\text{-}1 \qquad$ etc, (5.3)

3) Two more conditions are required on f''_i, and these are obtained by specifying the two end conditions, the conditions for x_1 and x_{n-1} points. In the algorithm developed for the purpose of this study, f''_1 and f''_{n-1} are both set equal to zero, giving the so-called natural cubic spline.

$$f''_1(x_1) = 0$$

$$f''_{n-1}(x_{n-1}) = 0 \qquad (5.4)$$

Specifying the end conditions leaves n-2 unknowns, which can be obtained by solving the n-2 equations.

The algorithm was developed in HiQ version 4.1. The user was asked to specify the parameters and stopping conditions of the cubic spline baseline fitting process. The program initially selected random nodes on the experimental spectrum, and a natural cubic spline fit was automatically produced according to those random interpolation points selected. The experimental spectrum was then compared with the fitted line produced. A fraction of the peak area corresponding to the fitted line was chopped out from all peaks (in our case, a 95% peak retaining was specified). A new set of random interpolation nodes (with more nodes than in the previous set) was selected by the program, and cubic spline fitting of the interpolation points was repeated. The new fit was compared to the area-reduced spectrum and peak area reduction was repeated. The

process was cycled until a user-specified stopping condition was reached, which is the variance calculated from the fitted line and the remaining peak area reduced spectrum. The program then subtracted the baseline from the experimental spectrum, producing a flattened baseline.

5.2.2 Resultant Spectra from Cubic Spline Baseline Correction

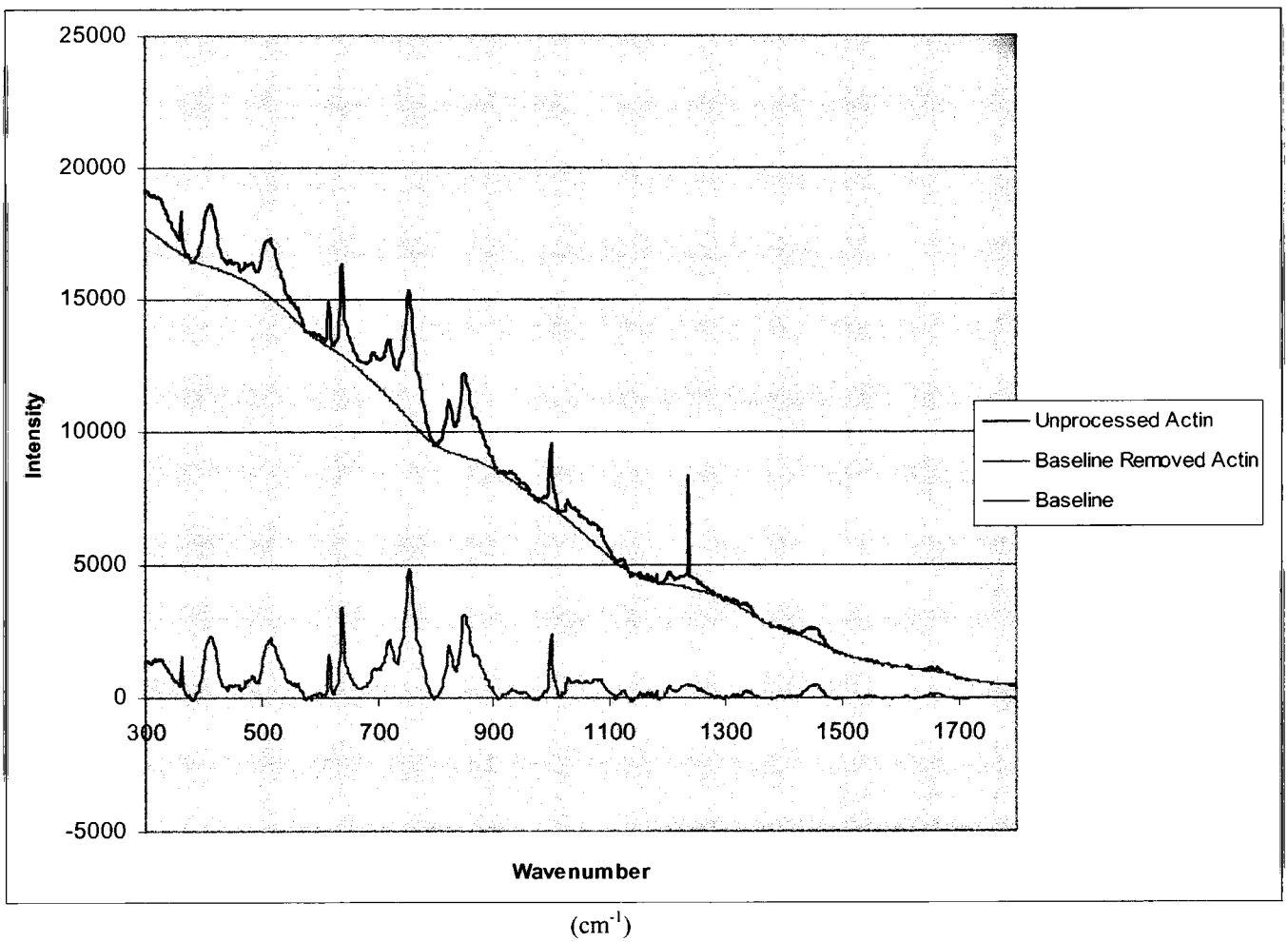


Figure 5.1 A cubic spline function was fitted to the unprocessed actin spectrum, the cubic spline fitted baseline is represented by the pink line, while the experimental actin spectrum is represented by the black line. The baseline removed actin spectrum, leaving spectral features, is represented by the red line.

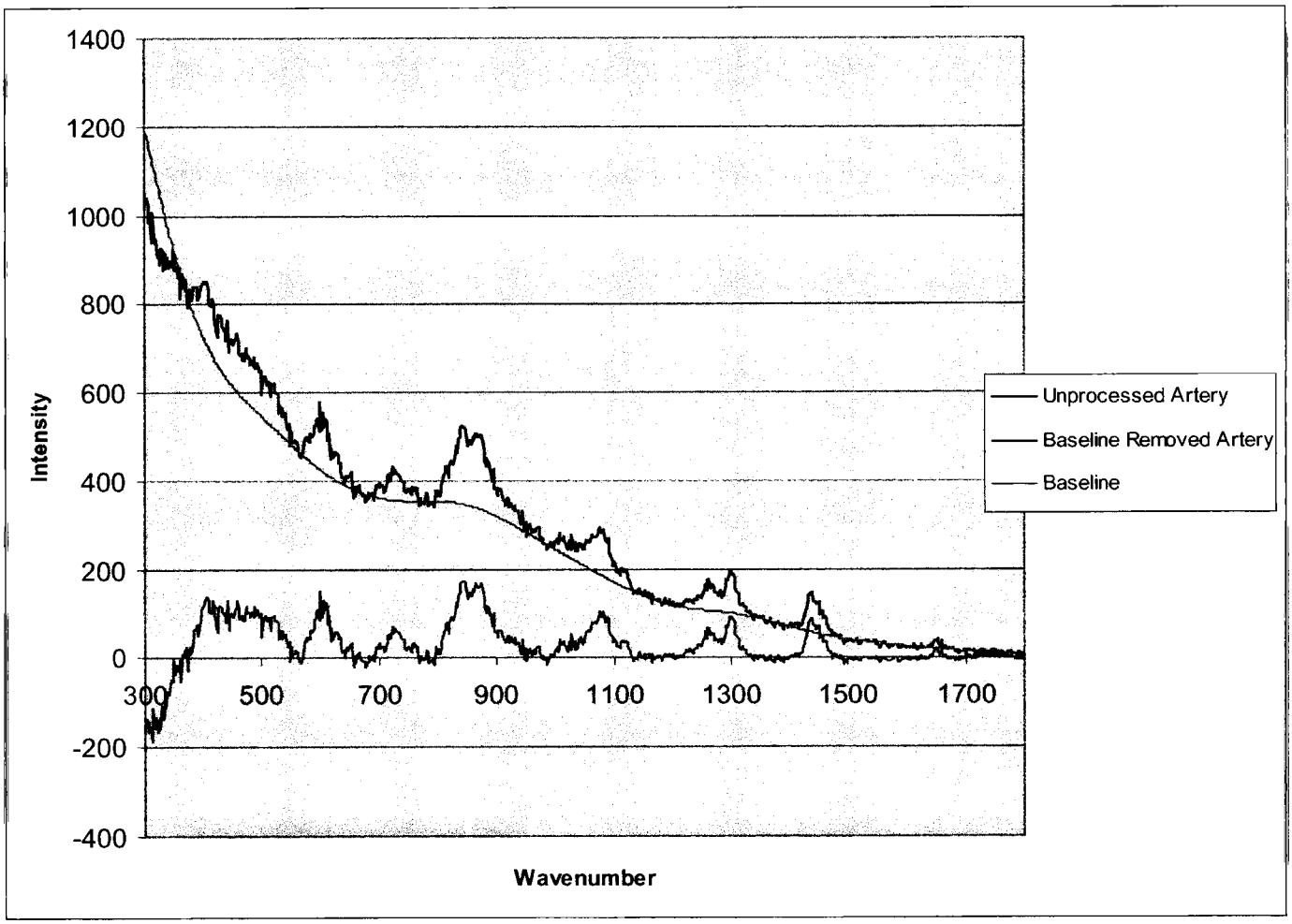


Figure 5.2 A cubic spline function was fitted to the unprocessed coronary artery spectrum, the cubic spline fitted baseline is represented by the pink line, while the experimental coronary artery spectrum is represented by the black line. The baseline removed coronary artery spectrum, leaving spectral features, is represented by the red line.

5.3 The Use of a Baseline Correction Method Formulated by Dietrich *et al.*

5.3.1 Theory of Dietrich's Method

Dietrich *et al.* formulated this method of baseline removal originally in 1991 [60]. The algorithm was employed in their study to exclusively remove baselines produced in 1D and 2D NMR spectra. However, since the method makes no assumption regarding the source or the functional form of the baseline distortions presented in different spectra, the approach was thought to also be applicable in other spectroscopic studies, such as Raman baseline removal.

To perform the baseline correction, the algorithm was again developed in HiQ version 4.1. The advantages of this peak-find algorithm are that it is fast, automatic and no manual parameter selection is needed.

The first step the program carried out was a moving average noise removal/smoothing of the experimental spectrum [60]. The user was asked to specify the number of data points that corresponded to the full width at half-maximum (FWHM) of an average peak, so that only noise but not peak was removed by this smoothing action. The next step the program performed was the calculation of the first derivative of the smoothed spectrum. This step resulted in a new spectrum, which was built by replacing every point in the spectrum by the difference between this point and the next point of the spectrum. A power spectrum was then calculated from the first derivative spectrum. These last two steps allowed baseline recognition to be performed more accurately, since the power spectrum of the first derivatived spectrum contained reduced baseline distortion, since baseline distortions were not amplified as much as spectral features by the squares of the signal intensity.

The estimation of the baseline required the finding of a correct threshold, which reliably differentiated the baseline from peak areas. To find the correct threshold automatically, the program first calculated the mean value and the standard deviation of the entire power spectrum. An initial threshold was set to mean + (3 x SDEV) [60]. A new MEAN and SDEV were calculated from the remaining points. The procedure was repeated automatically until no new point exceeded the final threshold of MEAN + (3 x SDEV).

All that remained was to determine whether a point belongs the to baseline or the peak. If only a single point exceeded the threshold and both of its neighbors did not, then this point belonged to the baseline. In turn, if both neighbors of a baseline point belonged to the peaks, this point was also called a peak.

The final step was to fit the recognized baseline points with a spline function mentioned in the earlier in the previous section. After the baseline was estimated, subtraction of the fit from the experimental spectrum yielded the baseline-flattened spectrum.

5.3.2 Resultant Spectra from Dietrich's Baseline Correction Method

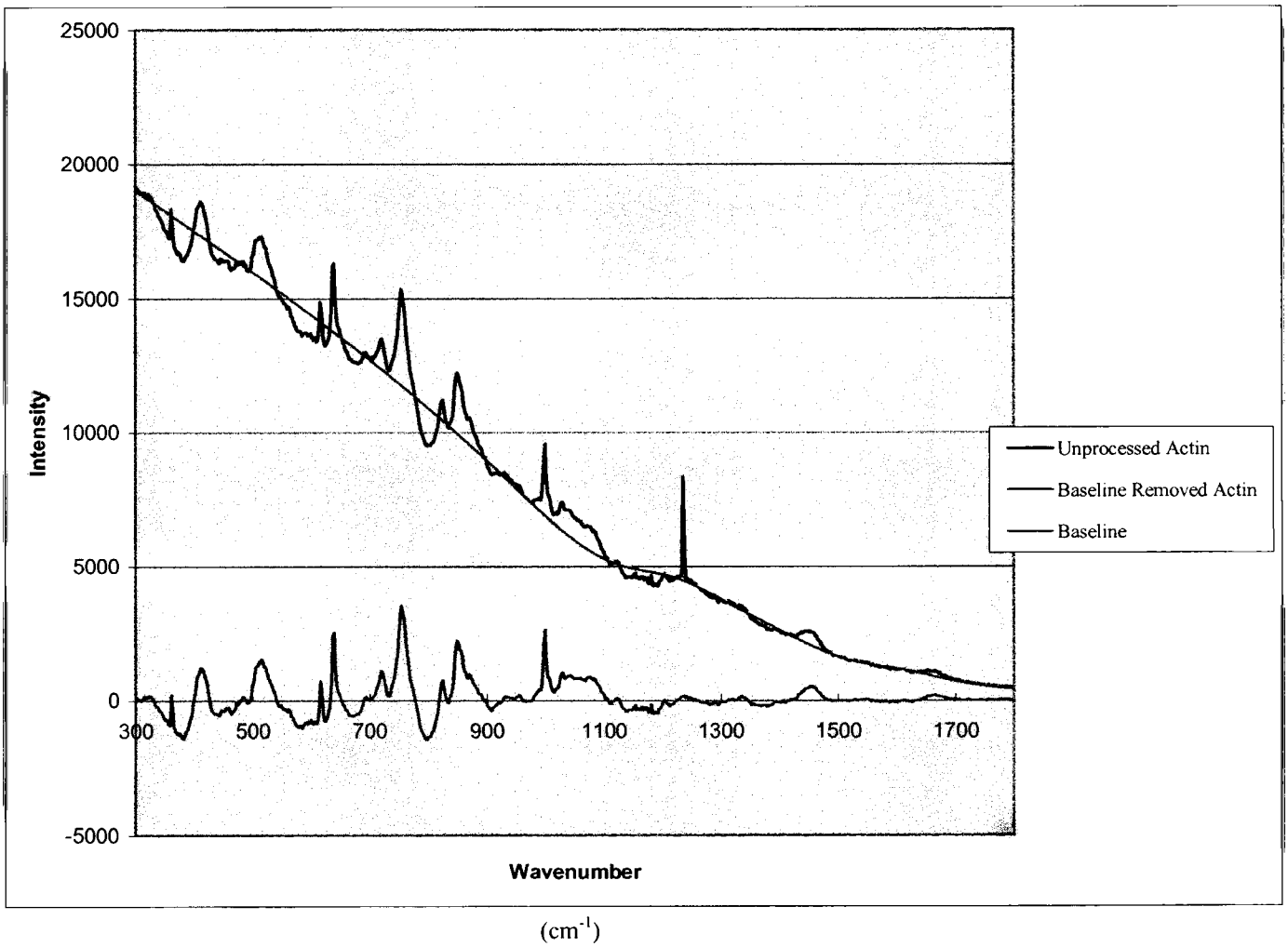


Figure 5.3 The same unprocessed actin spectrum as in figure 5.1 together with a baseline estimated by Dietrich's method from the unprocessed actin spectrum. The estimated baseline is represented by the pink line, while the experimental actin spectrum is represented by the black line. The baseline removed actin spectrum, leaving spectral features, is represented by the red line.

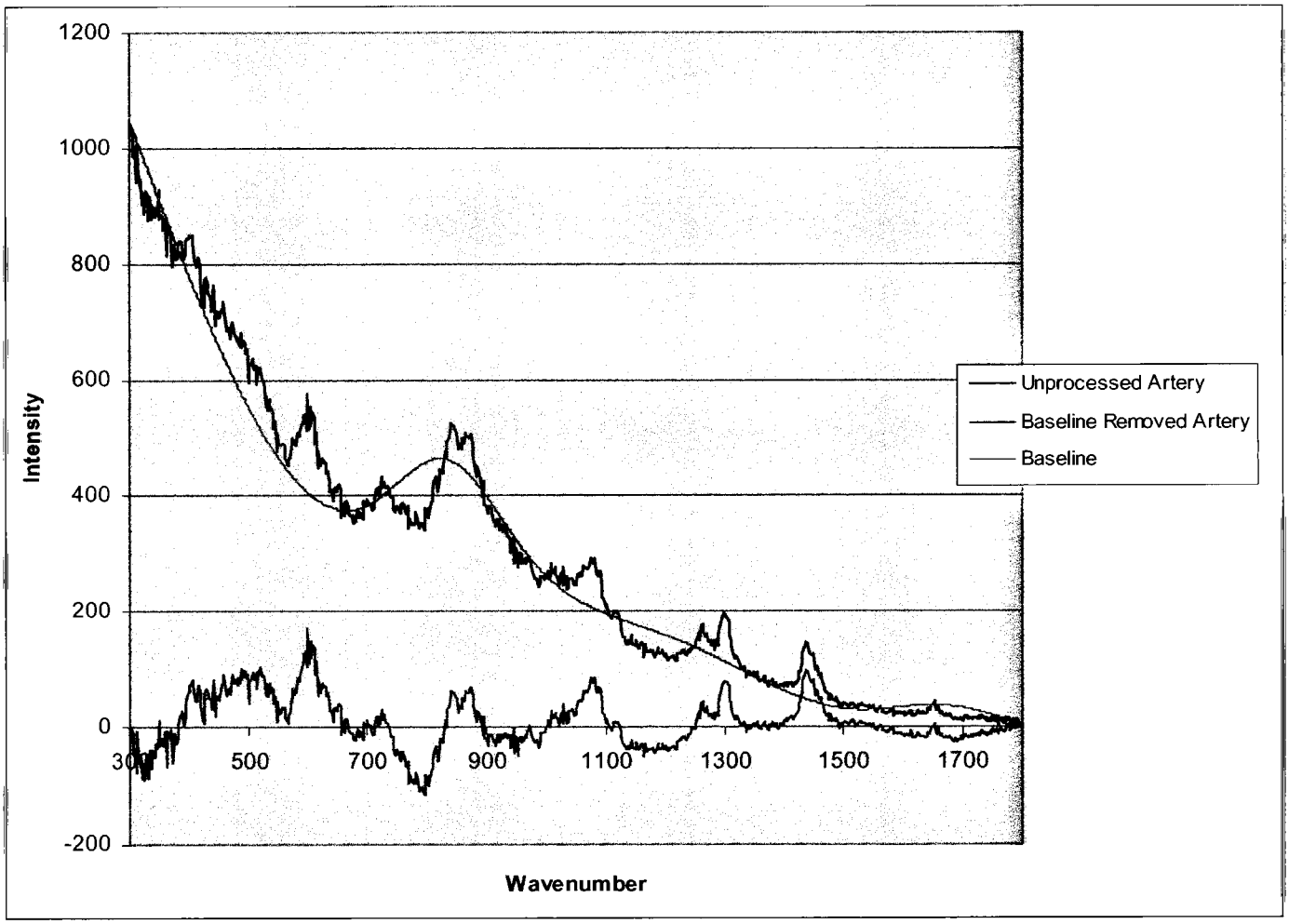


Figure 5.4 The same unprocessed coronary artery spectrum as in figure 5.2 together with a baseline estimated by Dietrich's method from the unprocessed coronary artery spectrum. The estimated baseline is represented by the pink line, while the experimental coronary artery spectrum is represented by the black line. The baseline removed coronary artery spectrum, leaving spectral features, is represented by the red line.

5.4 Comparison Between Cubic Spline Method and Dietrich's Method

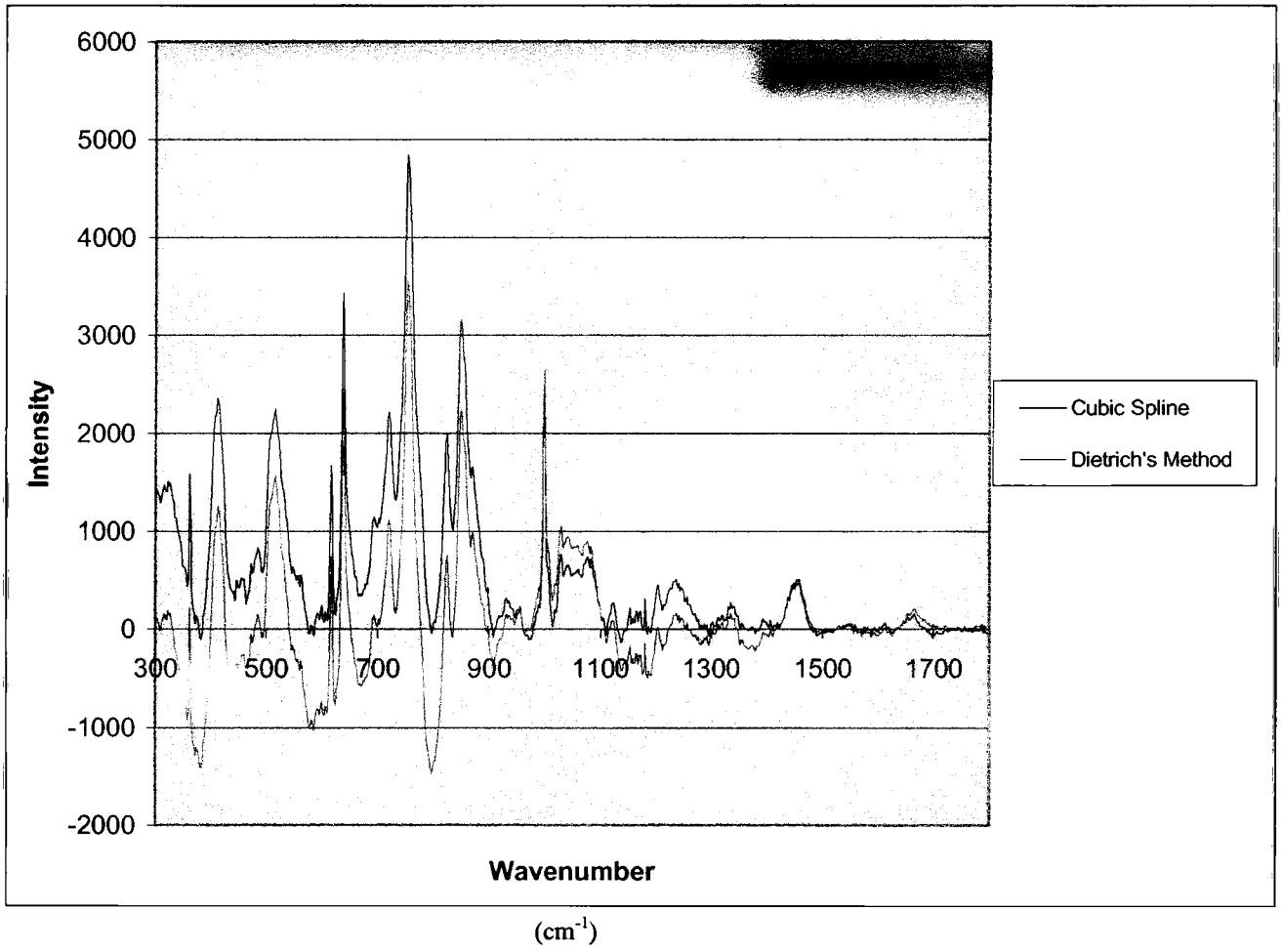


Figure 5.5. A comparison between the baseline corrected actin spectrum obtained by cubic spline function and the baseline corrected actin spectrum obtained by Dietrich's method. The resultant spectrum obtained by cubic spline function is represented with the black line, while the resultant spectrum obtained by Dietrich's method is represented with the red line.

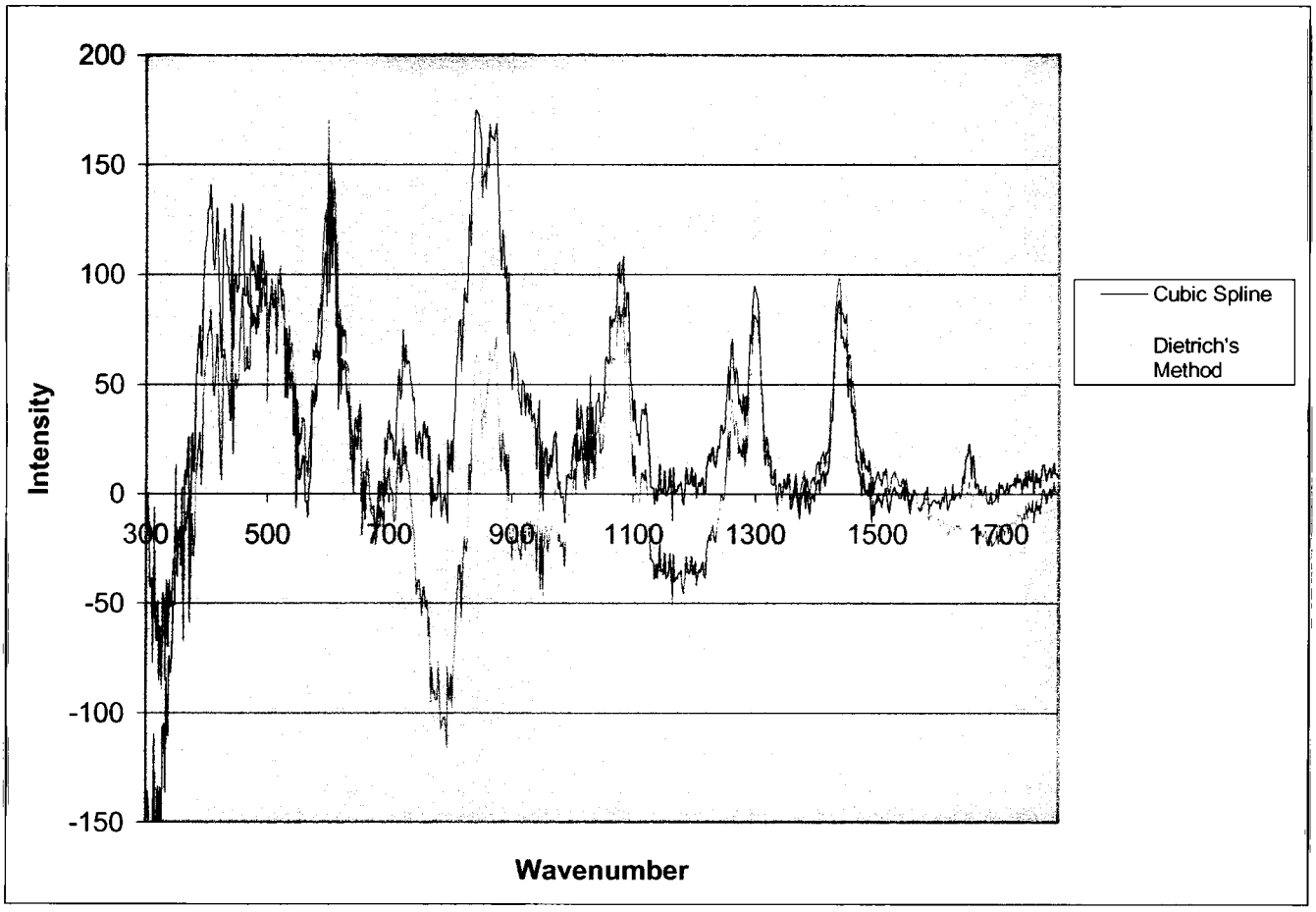


Figure 5.6 A comparison between the baseline removed coronary artery spectrum obtained by cubic spline function and the baseline removed coronary artery spectrum obtained by Dietrich's method. The resultant spectrum obtained by cubic spline function is represented with the black line, while the resultant spectrum obtained by Dietrich's method is represented with the red line.

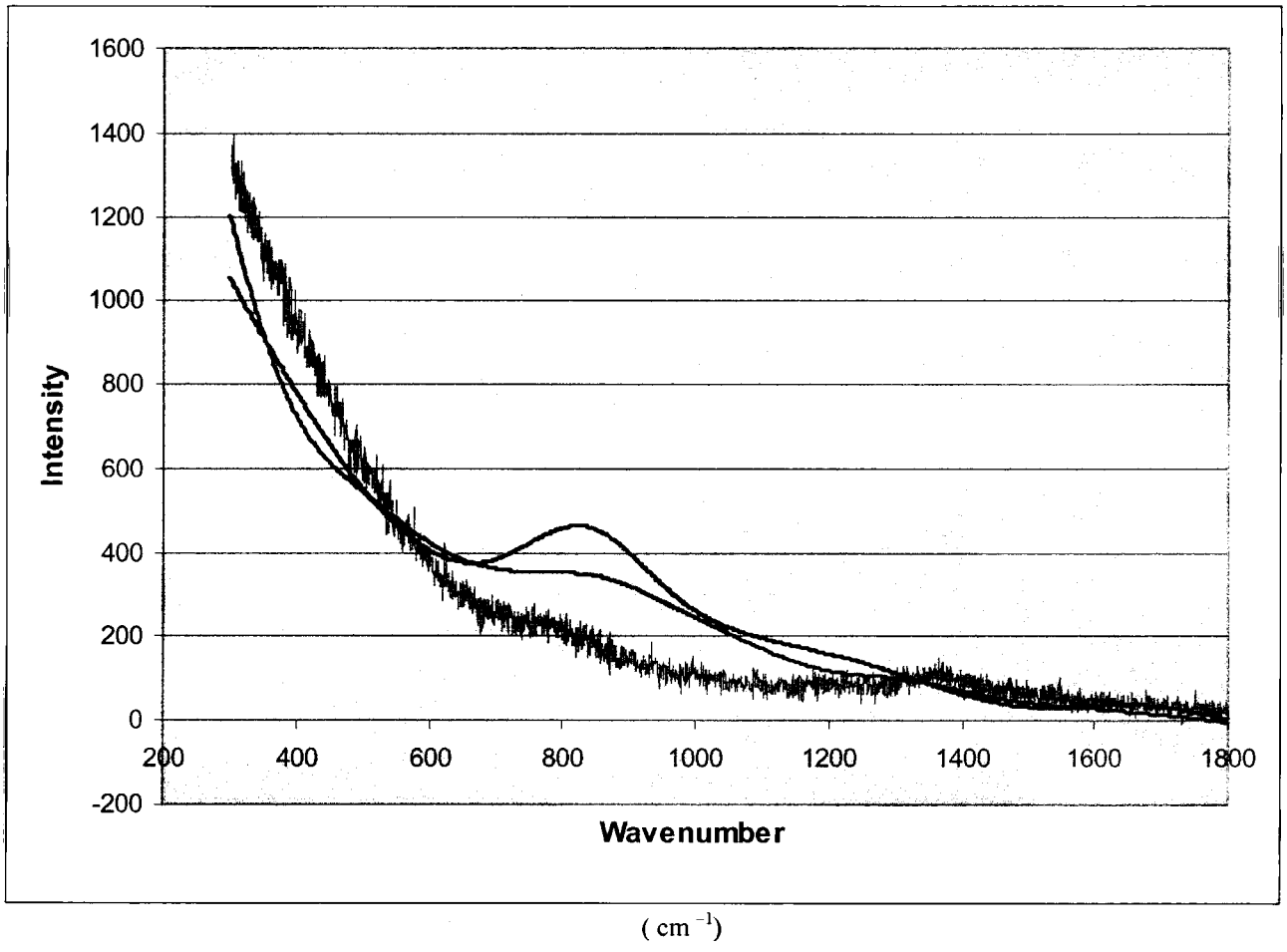


Figure 5.7 Comparison among the baseline of the coronary artery spectrum estimated by Dietrich's method, the cubic spline fitted baseline removed from the coronary artery spectrum, and the experimental Raman spectrum of OCT obtained under the same condition as the coronary artery spectrum. The estimated baseline obtained by spline function is represented by the red line, the estimated baseline obtained by Dietrich's method is represented by the green line, while the experimental background spectrum is represented by the blue line.

Using both cubic spline and Dietrich's baseline correction algorithms, a broad featureless sloping background due to OCT/optical elements was subtracted from the actin and artery spectra. The baseline corrected spectra of actin and artery results also looked very similar with sharp distinct peaks. The removed baseline employing both methods in the case of artery signal collection looked very similar to the experimental Raman spectrum of pure OCT on aluminium surface. However, upon a closer look (figure 5.7), it can be seen that the baseline removed by the cubic spline method seems to match the OCT spectrum better than the baseline resulted from Dietrich's method. The cubic spline method seemed to be more appealing also because it yielded less negative data points than in the case using Dietrich's method (figures 5.5 and 5.6). For instance, in the baseline corrected artery spectrum obtained by Dietrich's method, the baseline in the region from 300 –700 cm^{-1} was not determined well enough, and thus yielded a whole section of data below zero, and the new baseline was not flat either. This problem will affect the accuracy during normalization and any further analyses. For this reason, the cubic spline method was employed to correct the baselines of all experimental spectra in the study.

5.5 Normalization

Since each biochemical used in the basis set has its characteristic Raman scattering, and the actual concentration of these biochemicals in various locations of the arterial wall is not known, in order to accurately quantify the biochemical composition of the arterial wall, normalization of the all Raman spectra collected was necessary.

All artery Raman spectra, obtained after baseline correction, were normalized with respect to the maximum intensity of the spectral feature found between 1440-1455

cm^{-1} , which represent the C-C / C-H bending mode of most biochemicals composing the arterial wall [47].

The biochemical basis spectra, followed baseline removal, were also normalized with respect to same spectral feature located at 1440-1455 cm^{-1}, except for five of them. The Raman spectra of d-glucose and sucrose had the particular band corresponding to the C-C /C-H bending mode shifted up in wavenumber, were set to unity with respect to the spectral feature at 1465cm^{-1} instead. The Raman spectra of β-carotene, calcium carbonate, and calcium hydroxyapatite lacked C-C / C-H bending band; thus their spectra were normalized with respect to the spectral features at 1160cm^{-1}, 1090cm^{-1}, and 960cm^{-1} respectively [47].

Chapter 6

Generalized Two-Dimensional (2D) Correlation Analysis

6.1 Introduction - Generalized Two-Dimensional (2D) Correlation Analysis

6.1.1 Background

Generalized two-dimensional (2D) correlation spectroscopy is an extension of the original implementation of 2D correlation spectroscopy proposed by Noda in 1986 [66]. The original implementation of 2D correlation spectroscopy was only applicable to sinusoidal perturbations, while the generalized 2D method was revised by Noda to handle spectral fluctuations as an arbitrary function of any reasonable physical variables such as time, temperature, pressure, concentration and sample composition, etc [67]. The advantage of the generalized 2D Raman correlation is that it does not require any prior information of the sample [68].

The fundamental concept of the generalized two-dimensional (2D) correlation spectroscopy was first introduced during 1993 [66]. The generalized 2D correlation is a powerful, model-free spectroscopic tool to clarify and sort out complex variations by generating 2D correlation spectra useful for the determination of not-so-apparent correlated behavior of various spectral features. Because of the broad and robust applicability of 2D correlation analysis, this technique has received great attention in recent years, and has been successfully utilized in numerous areas of spectroscopy, mainly in NMR, but also including IR, Raman, fluorescence, NIR, X-ray, etc. Many applications of generalized 2D correlation spectroscopy have been reported. For instance, in protein research, generalized 2D correlation has been employed to study protein

secondary structure: denaturation, unfolding, and hydration [69]. In other studies, the technique has been used to probe molecular mechanisms during reactions [70,71].

In Raman spectroscopy, overlap of multiple species in a chemical system can create congested Raman spectra that are difficult to interpret. The extraction of component spectra is essential for the understanding of the structure and the dynamic of such multi-component systems. When the concentration of the chemicals is changing over a series of spectra, it is rather difficult to identify from such a simple stack of Raman spectra the subtle changes of spectral features, such as the development of shoulders and peak position shifts, and their complex relationships. Generalized 2D correlation spectroscopy has enabled the enhancement of the spectral resolution by resolving the correlations into real and imaginary components and by distributing the peaks over a second dimension [72]. The 2D asynchronous analysis, which will be discussed in later sections, is a particular powerful tool in analyzing the highly overlapped region of spectra.

6.1.2 Theory

The general principles of two-dimensional correlation analysis have been developed by Noda [72]. In this section, the mathematical approach for calculating the correlation intensities will be discussed.

Generalized 2D correlation analysis requires a series of perturbation-dependent spectra. In the present study, they are obtained by mapping across the cross section of the coronary artery samples.

From the location-dependent Raman scattering intensity I ($\Delta\nu$, x) observed during the mapping of the artery as a function of not only Raman shift $\Delta\nu$ but also the location,

x, one can obtain a mean-centered or an end-point centered dynamic spectral intensity variations,

$$\tilde{I}(\Delta\nu, x) = I(\Delta\nu, x) - \bar{I}(\Delta\nu) \qquad (6.1)$$

The dynamic intensity may be regarded as the difference or deviation spectrum of the individual location-dependent Raman spectrum I ($\Delta\nu$, x) with respect to the location-averaged (mean) reference spectrum, or the first or the last spectrum of the set (for either case, we call it $\bar{I}(\Delta\nu)$) obtained across the whole mapping distance X. (The continuous form of the equations are given.)

$$\bar{I}(\Delta\nu) = \int_0^X I(\Delta\nu, x)\, dx/X \qquad (6.2)$$

The synchronous and asynchronous 2D Raman correlation intensity, $\Phi(\Delta\nu_1, \Delta\nu_2)$ and $\Psi(\Delta\nu_1, \Delta\nu_2)$, which correspond to the real and imaginary parts of the complex cross-correlation function of the dynamic Raman spectral intensity variations, are given by

$$\Phi(\Delta\nu_1, \Delta\nu_2) = 1/X \cdot \int_0^X \tilde{I}(\Delta\nu_1, x) \cdot \tilde{I}(\Delta\nu_2, x)\, dx \qquad (6.3)$$

$$\Psi(\Delta\nu_1, \Delta\nu_2) = 1/X \cdot \int_0^X \tilde{I}(\Delta\nu_1, x) \cdot \hat{J}(\Delta\nu_2, x)\, dx \qquad (6.4)$$

where $\hat{J}(\Delta\nu_2, x)$ is the Hilbert transformation developed by Noda, for calculating the asynchronous component (phase shifted by $\lambda/2$ or 90°) without having to perform Fourier transform [73].

For the 2D correlation analysis, we used the same software (HiQ version 4.1, National Instruments Corp., Austin, TX), which was used in the baseline removal processing. The programming code is included in Appendix C. Applying the complex cross-correlation functions mentioned above, which yields real and imaginary components that contain information about the phase behavior of the variables.

The set of baseline corrected Raman data are then converted to the synchronous and asynchronous 2D Raman correlation spectra by using the relationships given in equations 6.1 to 6.4.

Normalization was not performed to the baseline corrected Raman spectra used in the 2D correlation study, because doing so would have taken away the correlation information at the normalized wavenumber (1440-1455 cm^{-1}). Assuming the excitation power of the laser was stable, the Raman intensity changes should solely be affected by the chemical composition difference occurring at different locations of the artery.

6.1.3. Meaning of Correlation Intensities/Plots

In 2D Raman correlation spectroscopy, the cross-correlation intensities of dynamic variations of Raman scattering signals, arising from the physiochemical changes imposed onto the system of interest, such as those created by mapping across the coronary artery samples, are plotted on two independent Raman shift axes.

The synchronous correlation intensity $\Phi(\Delta\nu_1, \Delta\nu_2)$ represents the coincidental or simultaneous changes of Raman spectral intensities measured at $\Delta\nu_1$ and $\Delta\nu_2$ during the mapping of the artery [72]. A synchronous 2D Raman spectrum is a symmetric spectrum with respect to a diagonal line corresponding to coordinates $\Delta\nu_1 = \Delta\nu_2$. Correlation peaks appear at both diagonal and off-diagonal positions.

Autopeaks located at positions of a synchronous 2D Raman spectrum originate from the correlation of each resolution element with itself. The magnitude of an autopeak, which corresponds to the autocorrelation function of spectral intensity variations, represents the overall change in the Raman scattering intensity of the sample observed at a given Raman shift during the mapping. The sign of autopeaks is always positive.

The peaks located at the off-diagonal positions of a synchronous 2D Raman spectrum are referred to as cross-peaks, which represent simultaneous or coincidental changes of spectral intensities observed at two different Raman shifts, $\Delta\nu_1$ and $\Delta\nu_2$ [72,74]. Such a synchronized change, in turn, suggests the possible existence of a coupled or related origin of the Raman intensity variations. The sign of cross-peaks can be positive or negative. The sign of a synchronous cross-peak becomes positive if the spectral intensities at the two Raman shifts corresponding to the coordinates of the cross-peak are either increasing or decreasing together as a function of location during mapping. On the other hand, a negatively signed cross-peak indicates one of the Raman intensities is increasing, while the other is decreasing.

An asynchronous 2D Raman spectrum $\Psi(\Delta\nu_1, \Delta\nu_2)$, or the imaginary part of the cross-correlation function, is antisymmetric with respect to the diagonal line. The asynchronous spectrum has no autopeaks and consists exclusively of cross-peaks located at off-diagonal positions. A positive peak for $(\Delta\nu_1, \Delta\nu_2)$ will have a negative peak for $(\Delta\nu_2, \Delta\nu_1)$. This antisymmetric appearance of the asynchronous spectrum is a result of the temporal information contained in the asynchronous component. An asynchronous cross-peak develops only if the intensities of two Raman intensities change out of phase (i.e.

delayed or accelerated) with each other [72,74]. This feature is especially useful in differentiating overlapped bands arising from Raman signals of different origins. For example, different spectral intensity contributions from individual components of a complex mixture may be effectively discriminated. Even if bands are located relatively close to each other, as long as the signature or the pattern of location-dependent variations of the Raman intensities are substantially different, asynchronous cross-peak will develop between their spectral coordinates. The sign of asynchronous 2D Raman cross-peaks can be either positive or negative. The sign of an asynchronous cross-peak becomes positive if the intensity change at $\Delta\nu_1$ occurs predominantly before $\Delta\nu_2$. It becomes negative, on the other hand, if the change occurs after $\Delta\nu_2$ [67].

6.2 2D Correlation Analysis of the Entire Arterial Wall.

In this section, rows of spectra were collected over the radii of artery specimens with different narrowing of the lumen, covering the three layers of the coronary artery (intima, media, and adventitia). The aim of employing 2D correlation analysis across different layers is to reveal spectral features (biochemicals) appearing and disappearing together in different arterial layers (by analyzing positive synchronous peaks). If the peaks are not appearing together, the order of their appearance along the radius of the coronary artery can also be revealed (by analyzing negative synchronous peaks and asynchronous plots). Hence, from the different types of correlation spectra, the dynamic behavior of the peaks can be deduced.

In addition, the synchronous spectrum contains information about features that change in-phase, while the asynchronous spectrum displays information regarding out-of-

phase covariance. The synchronous and asynchronous spectra provide complementary information. The presence of a synchronous peak and the absence of an asynchronous peak indicate peaks that arise from the same biochemical (perfectly correlated in-phase). Conversely, the presence of an asynchronous peak and the absence of a synchronous peak indicate peaks that arise from the two different constituents (perfectly correlated out-of-phase).

Assignments of the NIR Raman bands of coronary arteries are listed in Table 6.1. The assignments were based upon previous research [45, 47]. The Raman bands below 650cm^{-1} were not included because they were not assigned in previous studies.

Table 6.1. Assignments of the bands in the NIR Raman spectra of coronary artery.

Frequency (cm^{-1})	Assignment	Present in
650	sterol ring vibration	free cholesterol and cholesterol ester
702	steroid nucleus	free cholesterol and cholesterol ester
719	choline (sym) stretching	phospholipids
762	O-P-O (sym) stretching	phospholipids
850	C-C or C-H bending	proteins, lipids, triglycerides, phospholipid and cholesterol
855	C-C stretching of proline/hydroxyproline	collagen
870	O-P-O (asym) stretching	phospholipids
870	steroid nucleus	free cholesterol and cholesterol ester
923	steroid nucleus	free cholesterol and cholesterol ester
933	C-C stretching of proline/hydroxyproline	collagen
940	C-C or C-H bending	proteins, lipids, triglycerides, phospholipid and cholesterol
957	steroid nucleus	free cholesterol and cholesterol ester
961	PO_4^{3-} (sym) stretching	calcium hydroxyapatite
1004	phenylalanine	proteins
1026	C-C or C-H bending	cholesterol ester
1034	C-C or C-H bending	proteins, lipids, triglycerides, phospholipid
1058	C-C or C-H bending	free cholesterol
1080-1090	CO_3^{2-} (sym)	calcium carbonate
1085	C-C stretching	proteins, lipids, triglycerides, phospholipid and cholesterol
1104	desmosine/ isodesmosine	elastin
1130	sterol ring vibration	free cholesterol and cholesterol ester
1159-1160	C=C stretching	β-carotene
1176	C-C stretching	proteins, lipids, triglycerides, phospholipid and cholesterol
1264-1268	amide III vibration	proteins
1270-1274	C-C or C-H bending	triglycerides
1300	C-C or C-H bending	free cholesterol
1336	desmosine/ isodesmosine	elastin
1336	C-C or C-H bending	proteins and lipids
1440-1455	C-C or C-H bending	proteins, lipids, triglycerides, phospholipid and cholesterol
1523	C-C stretching	β-carotene
1654-1671	C=C stretching	triglycerides
1660-1664	amide I vibration	proteins
1667-1668	C=C stretching	cholesterol crystal
1735-37	ester linkage	cholesterol ester
1747	ester	triglycerides

6.2.1. Completely (100%) Occluded Coronary Artery

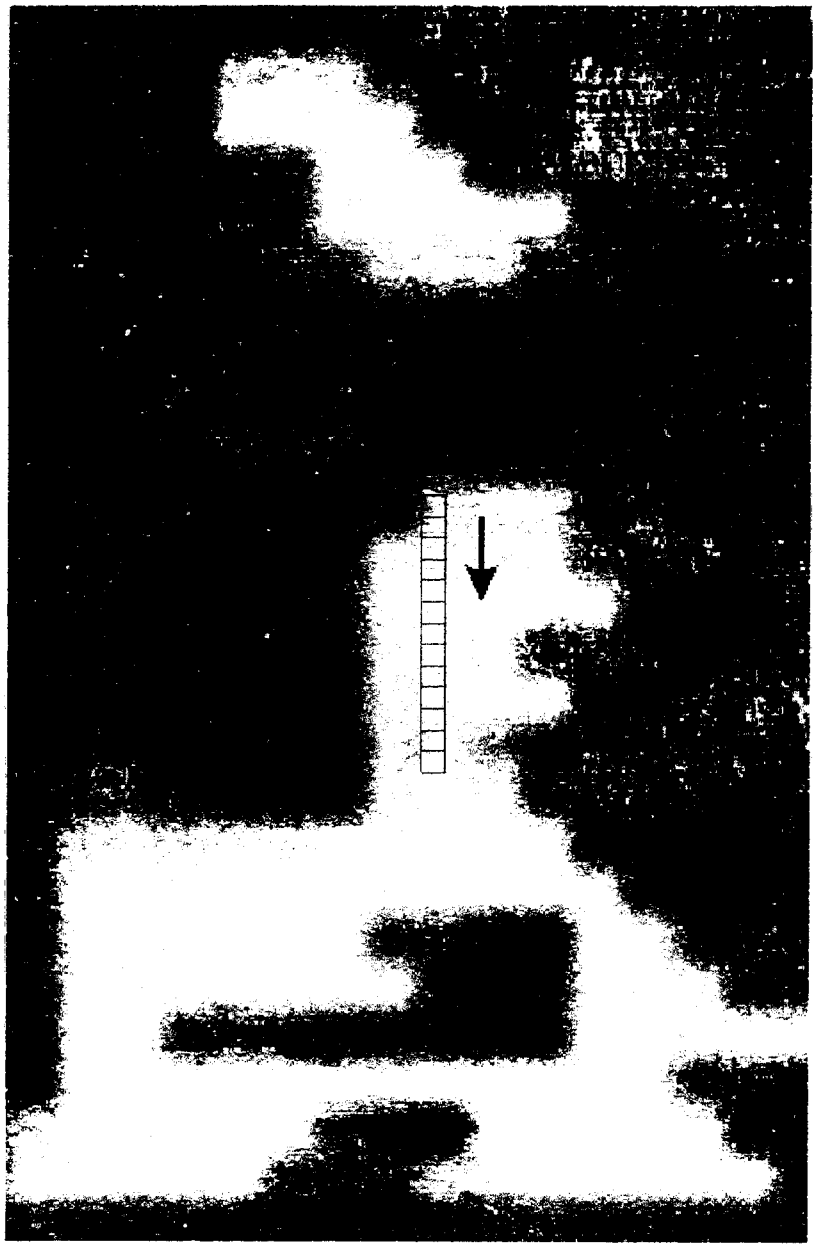
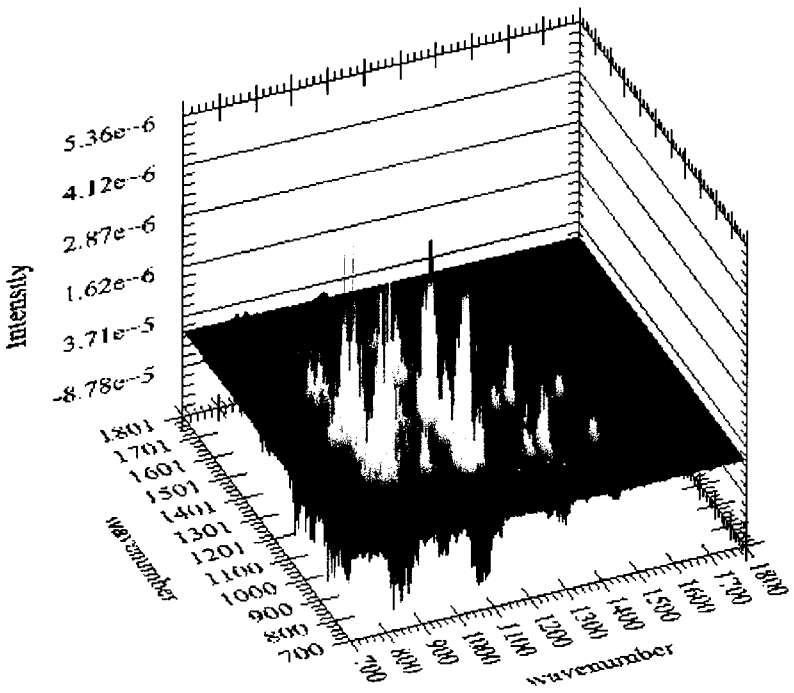


Figure 6.1 Photo of a cross section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 100%. Spectra were collected at the center of each grid, starting from the adventitia to the intima (direction indicated by the arrow) in 100μm-step.

a)



b)

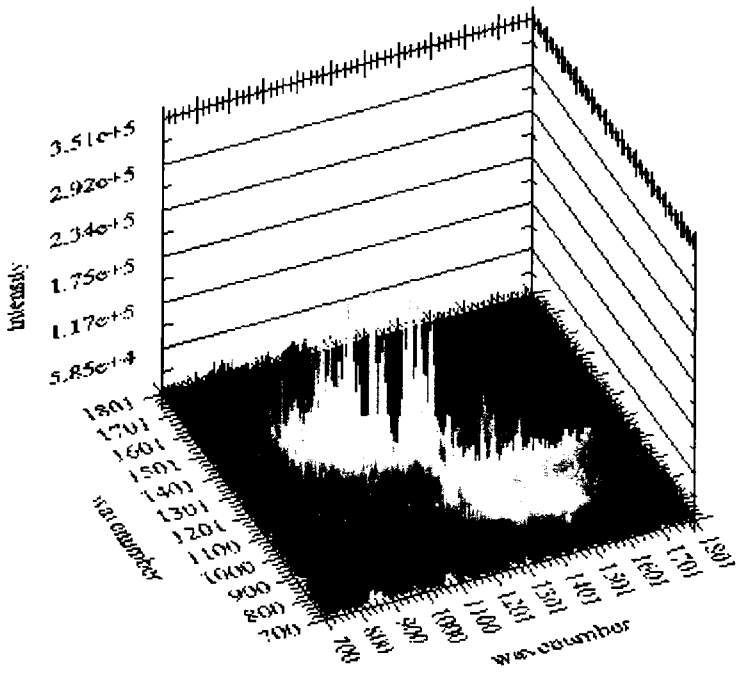
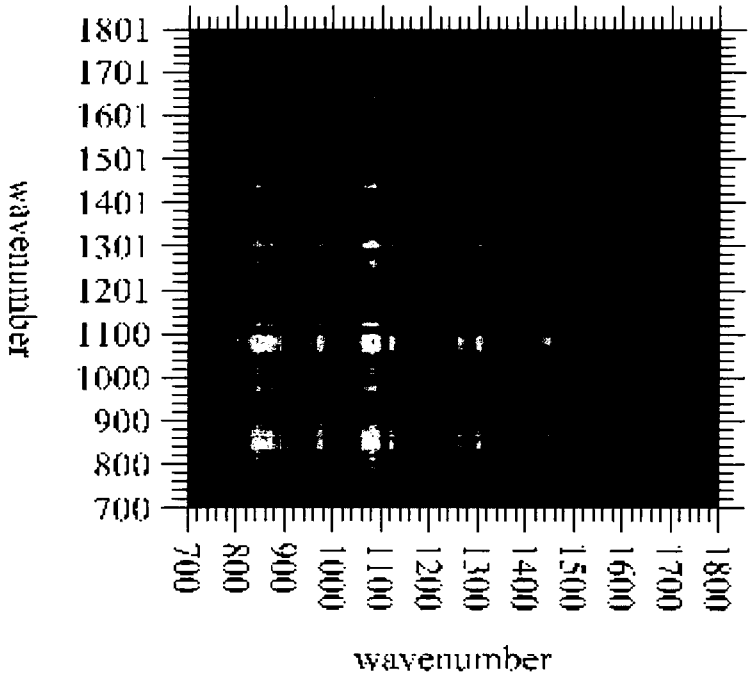


Figure 6.2 (a) A 3-D representation of the synchronous 2D Raman correlation spectrum of the mapping of the 20μm section of the 100% narrowed coronary artery in the spectral region from 700-1800 cm^{-1}. Purple areas represent the negative peaks. (b) A 3-D representation of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. All peaks are positively shown.

a)



b)

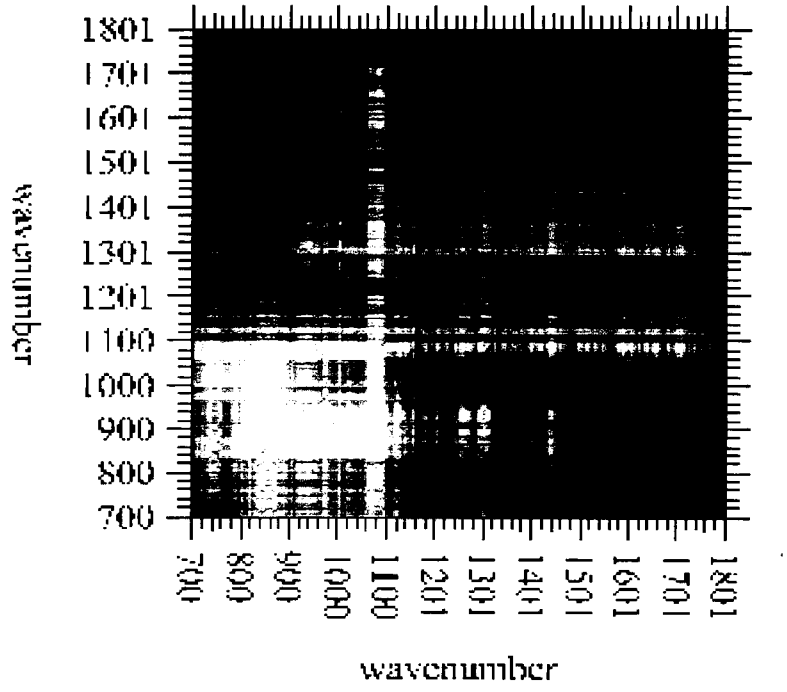


Figure 6.3 (a) A 2D contour plot of the synchronous 2D Raman correlation spectrum of the 20μm section of the 100% narrowed coronary artery in the spectral region from 700-1800 cm^{-1}. Purple areas represent the negative peaks. (b) A 2D contour plot of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. All peaks are positively shown.

Either a 3D mesh diagram (figure 6.2) or a 2D contour plot (figure 6.3) of the synchronous and asynchronous spectra can be used to visualize the results. The intensity of the correlations and the sign of the peaks are more easily viewed using a 3D diagram. However, the 3D presentation may hide some of the peaks and comparison of peak heights may be difficult. Due to these problems, the correlation spectra are typically displayed as contour plots in literature [74]. The 2D contour plots have the advantage of ease of comparison.

6.2.1.1 Positive Synchronous Peaks

Numerous autopeaks and cross-peaks are observed in figure 6.3 (a), indicating the presence of Raman spectral intensity changes induced by location changes. Positive cross-peaks, which correspond to 850/1085 cm^{-1} has the highest intensity. This result is consistent with the fact that these two bands represent the C-C / C-H stretching/bending mode of most biochemicals (fats, lipids, cholesterol, proteins) composing the arterial wall (bending mode: 850cm^{-1}, stretching: 1085cm^{-1}). The intensity corresponds to 870/1085 cm^{-1} is strong as well. The phospholipid's O-P-O stretching band happened at the exact same position to cholesterol's steroid band, and thus 1085cm^{-1} (which belongs to both cholesterol and phospholipid) shows a particular strong positive correlation with 870 cm^{-1}. Medium intensity positive peaks were noticeable from the figure at various locations: 850/960, 870/960, 850/1130, 870/1130, 850/1300, 870/1300, 850/1440, 870/1440, 960/1085, 960/1130, 960/1300, 960/1440, 1085/1130, 1085/1300, 1085/1440, 1130/1300, 1130/1440, 1300/1440. This is reasonable, because all these five bands are intense bands found in cholesterol, cholesterol esters, lipids, phospholipids, and triglycerides for example. Correlation at around 1176cm^{-1} (C-C stretch in cholesterol,

cholesterol esters, lipids, phospholipids, and triglycerides) with the signals mentioned above should be seen. However, its absence might due to its Raman signal being too low and intensity change not significantly detectable. Other positive cross-peaks can be seen from the diagram, what they indicate, however is that the signals at the corresponding wavenumbers either increasing or decreasing together, and are peaks that show up together in the Raman spectrum of a single chemical. Those peaks can more easily be seen from the 3D mesh diagram.

Positive cross peaks could have been mis-generated due to the unstable laser power. For instance, if the laser power drops, all Raman intensities drops, and thus all peaks are being interpreted as correlating with each other [75]. This could have generated small correlation peaks making the plot look extremely noisy.

6.2.1.2 Negative Synchronous Peaks

Numerous negative cross peaks can be seen from the same figure (figure 6.3a), and they represent those Raman signals that uncorrelated with each other. Even though the intensities of the negative peaks are lower compared to those of the positive peaks as shown in figure 6.2(a), the intensity unit is arbitrary and not related in any manner. The most easily noticeable negative cross peaks correspond to 760/855, 760/870 and 760/1080. A negative correlation (i.e. not increasing nor decreasing together) was observed for phospholipids (represented by the O-P-O stretching at $760cm^{-1}$) and collagen (represented by $855cm^{-1}$ band). This was verified by chemical composition analysis (figure 7.5) in the following chapter. A negative correlation was also observed between phospholipids ($760cm^{-1}$) and cholesterol (represented by the $870cm^{-1}$ band), and between phospholipids ($760cm^{-1}$) and calcium carbonate (represented by the $1080cm^{-1}$

band). Phospholipids can be found in the arterial adventitia while cholesterol and calcium carbonate are more abundant in the arterial media and intima, hence, the discorrelation between phospholipids and the two constituents was observed. The 2D synchronous correlation analysis has shown the ability to separate and detect bands [75] that are not identifiable from the 1-D Raman spectra when there is more than one component contributing to the same Raman peak.

Many other negative cross peaks can be observed. However, some of them are hard to identify because the intensities shown are so noisy. In addition, a peak shift in wavenumber might have occurred due to variation in the chemical make up at different locations. The peak shifts have resulted in negative cross peaks appearing next to the positive peaks, interfering the interpretation of the negative cross peak information obtained [76].

6.2.1.3 Asynchronous Peaks

The asynchronous 2D Raman correlation spectrum shown in Figure 6.3(b) can reveal even more intriguing results. The order of the appearance of different Raman signals (hence, the order of appearance of biochemicals) over the various layers of the coronary artery can be disclosed. However, because a programming code error occurred in the asynchronous intensity calculation, all intensities were plotted as positive. The problem was not fixed by the time this thesis was submitted. Nonetheless, the asynchronous correlation diagrams were worth discussing. The negative correlation between phospholipids and collagen mentioned above can further be explored in the asynchronous plot. Even though not very intense, two peaks can be observed at 760/855 and 855/870. The band at 760 and 870 cm^{-1} belong to phospholipids while the band at

855cm^{-1} belongs to collagen. Since all asynchronous peaks appeared positive, the order of appearance of the two constituents can be determined. However, further chemical composition was performed (figure 7.5) and will be discussed in chapter 7, and from the results there, we know that phospholipids appeared in the adventitia first and its concentration dropped in the media where collagen started to appear. Hence, 760/855 and 870/855 intensities should be positively expressed while 855/760 and 855/870 should be shown in the negative axis. The algorithm will be corrected and the order of the appearance of the Raman signals (hence the order of the appearance of the constituents) can be determined.

Peak shifts in wavenumber due to variation in the compositional make up at different locations have resulted in extra asynchronous peaks. Their presence interferes with the interpretation of the information obtained. However, if one is interested in knowing how peak shifts in various different compositions, these extra peaks may become very informative.

6.2.1.4 Peak Shape

Finally, it might not be noticed from the 2D correlation plots, but it is worth to mention that correlation peaks are Gaussian in shape because the spectral peak intensities vary in a Gaussian manner [67].

6.2.2. 0-25% Occluded Coronary Artery

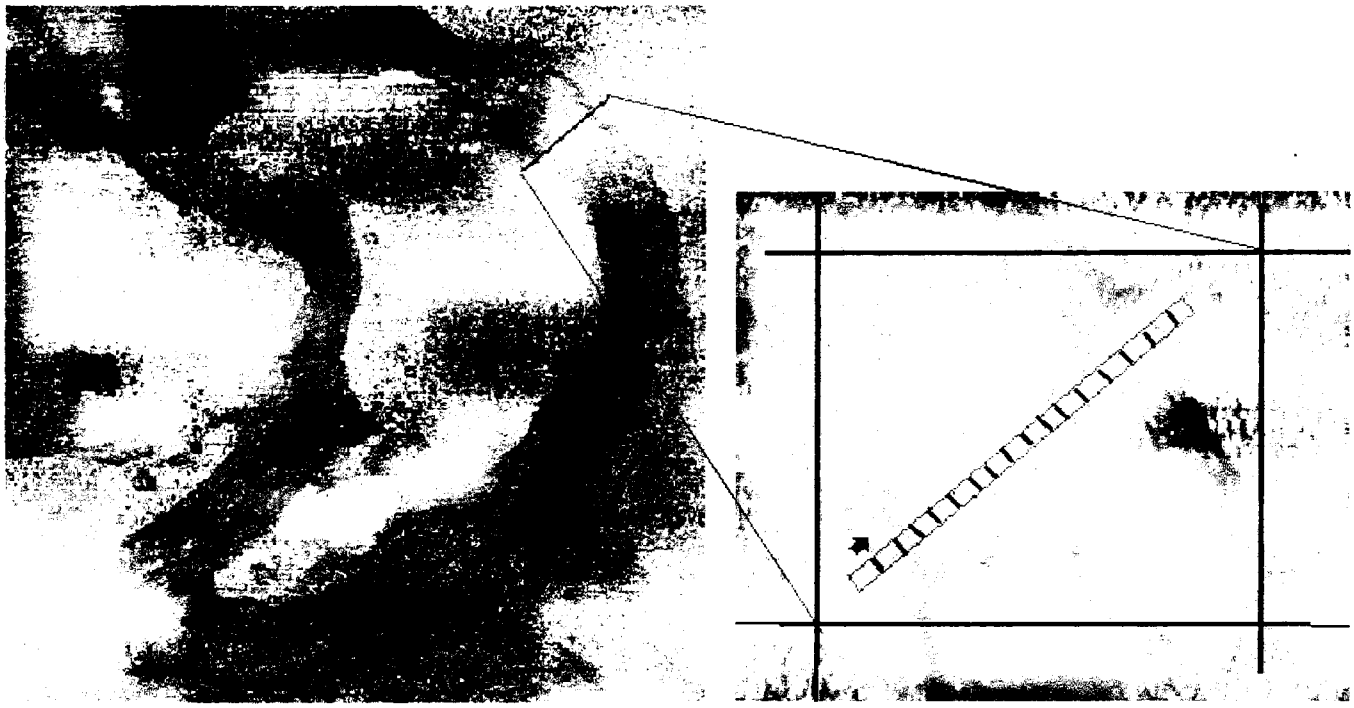
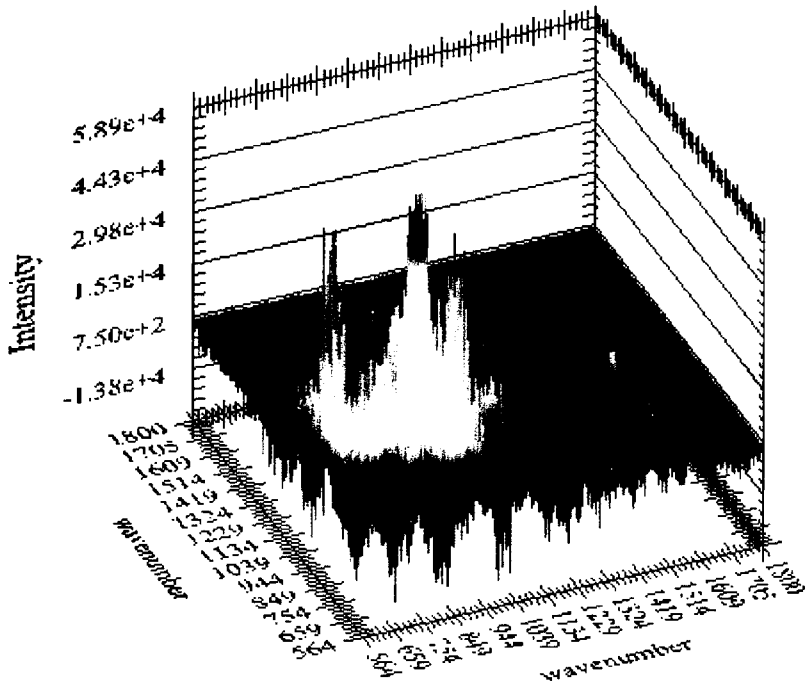


Figure 6.4 Photo of a cross section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 0-25%. Spectra were collected at the center of each grid, starting from the intima to the adventitia (direction indicated by the arrow) in 35μm-step.

a)



b)

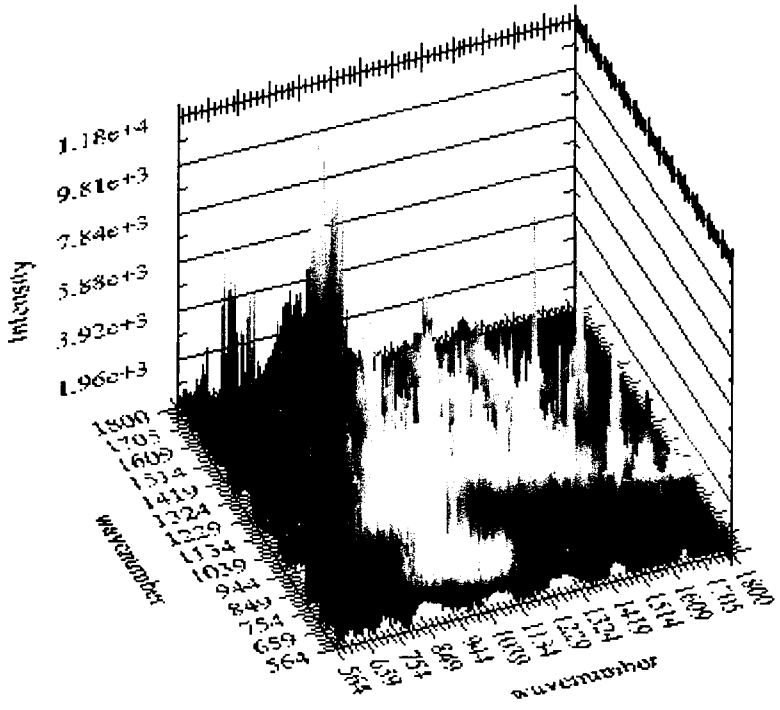
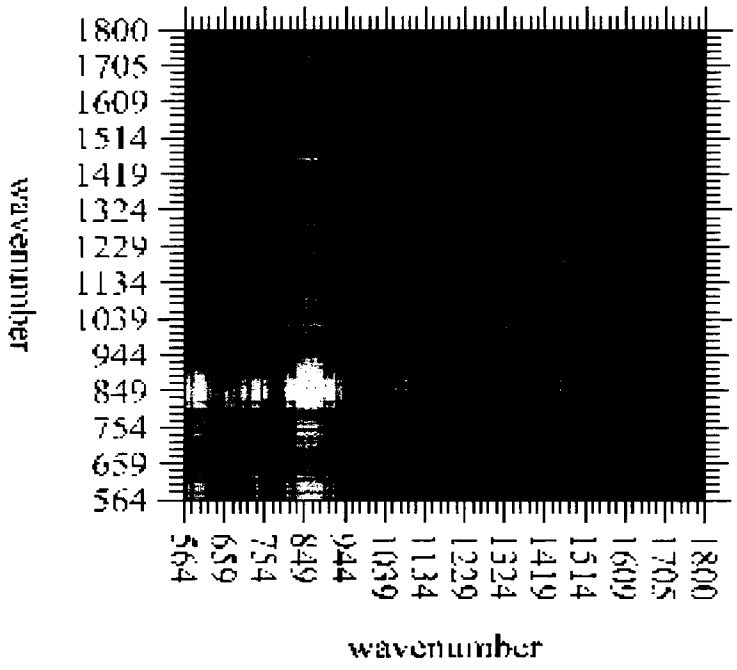


Figure 6.5 (a) A 3-D representation of the synchronous 2D Raman correlation spectrum of the mapping of the 20μm section of the 0-25% narrowed coronary artery in the spectral region from 560-1800 cm^{-1}. Purple areas represent the negative peaks. (b) A 3-D representation of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. All peaks are positively shown.

a)



b)

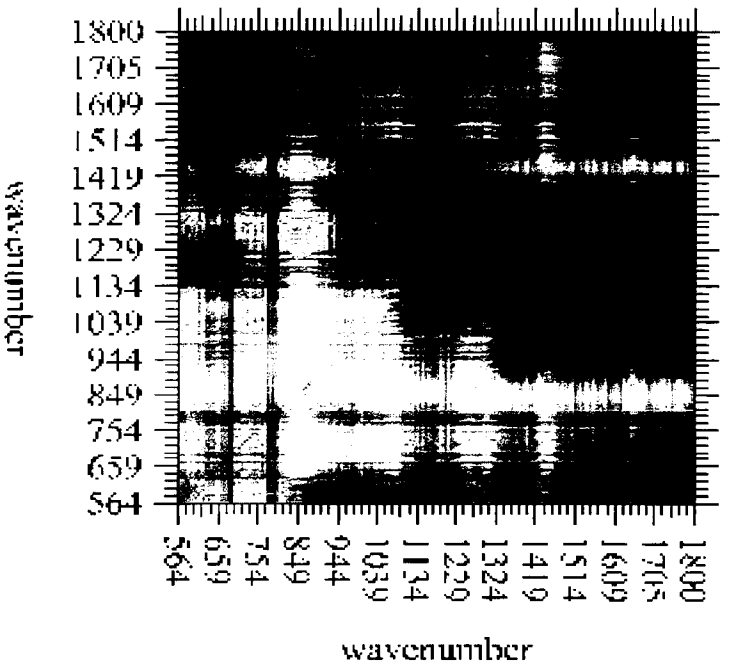


Figure 6.6 (a) A 2D contour plot of the synchronous 2D Raman correlation spectrum of the mapping of the 20μm section of the 0-25% narrowed coronary artery in the spectral region from 560-1800 cm^{-1}. Purple areas represent the negative peaks. (b) A 2D contour plot of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. Light areas represent the positive peaks, while dark areas represent the negative peaks. All peaks are positively shown.

6.2.2.1 Non-smoothed Raman Spectra

Again, numerous autopeaks and cross-peaks are observed in figure 6.5 (a), indicating the presence of spectral intensity changes induced by location moves. The result of the plot was very similar to the one in figure 6.3 (a). Even though there are many spurious peaks appearing, with effort, one can identify the barely visible signal peaks from the noise peaks in the plots. The locations of the positive cross-peaks and negative cross peaks are almost identical to those in the earlier example. The same thing applies to the asynchronous plot (figure 6.5(b)) in that similar intensities can be observed at about the same locations of the plot as in figure 6.3 (b). Thus, the interpretation of individual peaks is not repeated in this section.

What should be noted though is that the 2D correlation of the artery in the spectral region from 560 – 700 cm^{-1}, which was not performed in the earlier example, showed quite a lot of correlation and discorrelation. The Raman signatures within 560 – 700 cm^{-1} are indeed very intense. Since their assignments are very complicated, they were ignored in this study. However, these Raman signatures can be very informative if future research can pay more attention and more effort studying the Raman properties within this spectral range.

Another interesting fact is that the peaks are obscured in the plots due to extremely noisy results. It might because Raman spectra were collected in 35μm step as compared to 100μm step in the previous example. The previous sample was a 100% narrowed artery, which was mostly composed of lipid-rich plaque. In the current example, finer steps were performed and variation within different layers of the arterial wall was detected.

6.2.2.2 Smoothed Raman Spectra

For this reason, smoothing was performed on the baseline subtracted Raman spectra to remove unwanted noise. The smoothing function employed was a moving average filtering technique. Since the basic theory of a moving average filter has been covered extensively elsewhere [60,77], its description is not mentioned here. The window size employed to smooth the Raman spectra of the artery was nine points. The smoothing was performed using HiQ version 4.1, and the program code is included in Appendix D. After application of moving average filter to the artery Raman spectra, the resulted peaks in the 2D correlation plots (figure 6.7) now became apparent. Spurious peaks disappeared, and one can identify the signal peaks in the synchronous/asynchronous contour plots (figure 6.8) more easily.

However, the filtering or smoothing process might have thrown away important information especially when Raman signals are generally weak and background is extremely noisy. Thus, there is a trade-off between easier interpretation of 2D correlation plots versus the amount of information lost due to smoothing process, and further decision will be needed on this concern.

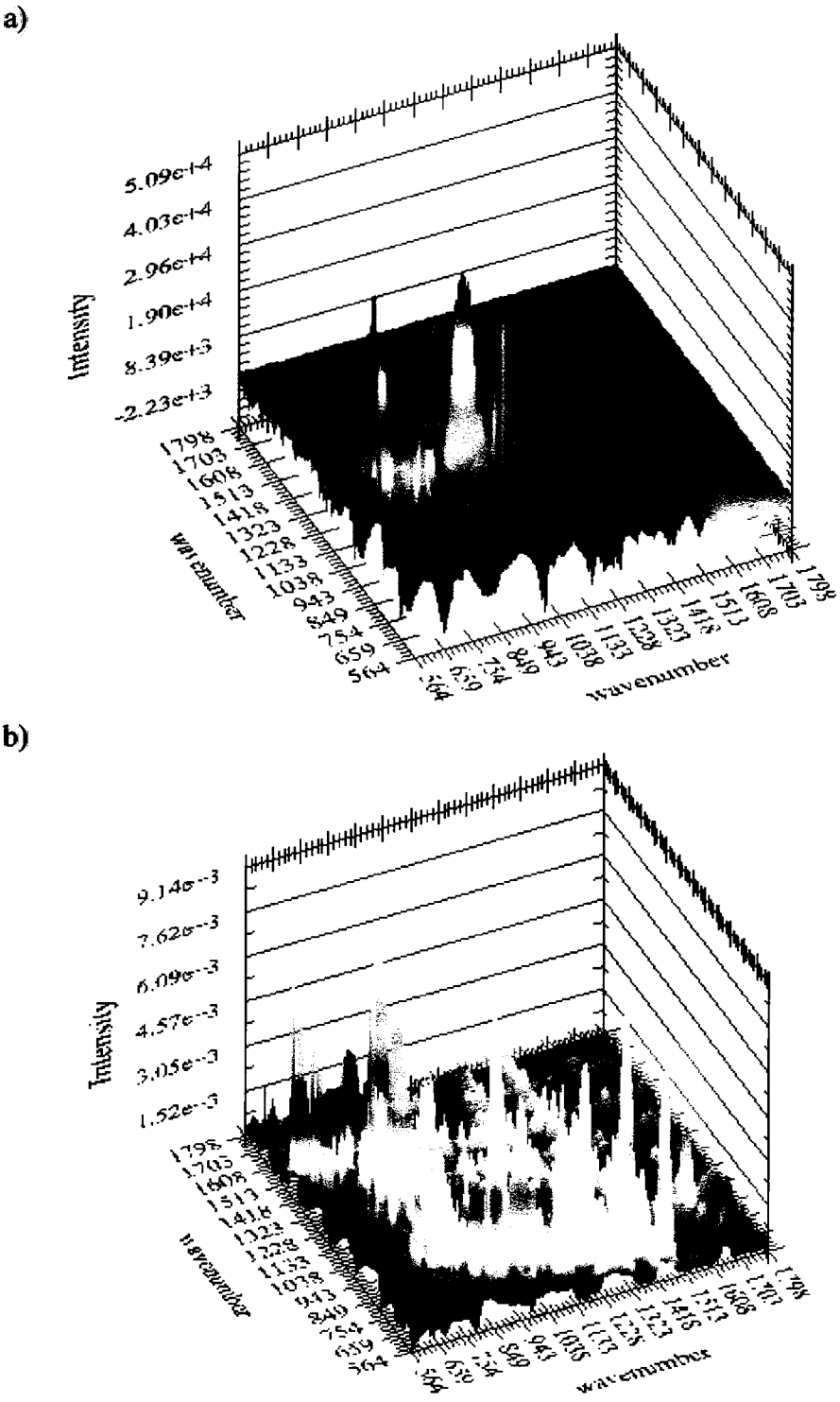


Figure 6.7 (a) A 3-D representation of the synchronous 2D Raman correlation spectrum of the mapping of the 20μm section of the 0-25% narrowed coronary artery in the spectral region from 560-1800 cm^{-1}. The Raman spectra were smoothed with a moving average filter after baseline correction. Dark purple areas represent the negative peaks. (b) 3-D representation of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. All peaks are positively shown.

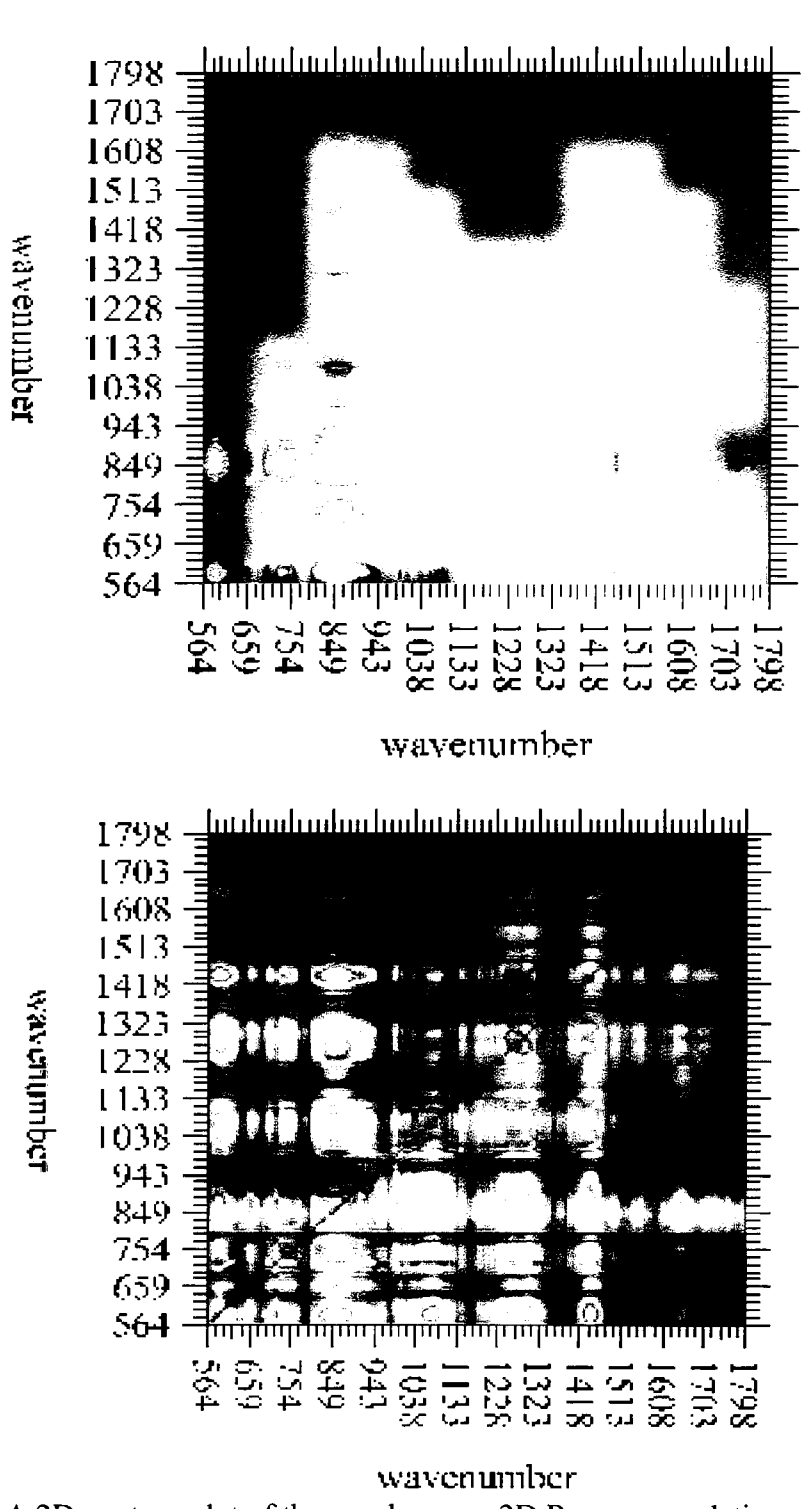


Figure 6.8 (a) A 2D contour plot of the synchronous 2D Raman correlation spectrum of the mapping of the 20μm section of the 0-25% narrowed coronary artery in the spectral region from 560-1800 cm^{-1}. The Raman spectra were smoothed with a moving average filter after baseline correction. Dark purple areas represent the negative peaks. (b) A 2D contour plot of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. All peaks are positively shown.

6.2.2.3 Complement Chemical Composition Analysis

The power of generalized 2D correlation analysis is its ability to apply to any number of spectra. The asynchronous spectra is useful to reveal information on the order of the appearance of different Raman signals, thus the order of appearance of the constituents over the various layers of the wall of coronary artery.

Across the different layers of the arterial wall, multiple analytes are varying with respect to sample location, and the analyte concentrations may vary at different rates. In addition, a biochemical component may be present in one layer, absence in the other layer, and then appear in the other layer again (e.g. phospholipids). When a plaque is formed, the chemical structure within the plaque is even more complicated. Hence, the whole set of spectra collected along the radius can be subdivided into many sections (a few spectra each), each with a very narrow distance, and 2D correlation analysis be performed to each subsection. In such a case, more information about the biochemical relationship in each individual layer can be determined. 2D correlation analysis can provide useful qualitative information about the order of the appearance of the constituents that complement the results of the chemical composition analysis.

6.3. 2D Correlation Analysis of the Arterial Intima/Plaque

In this section, Raman spectra were collected along a small section of the arterial intima. Intima is the focus for the two reasons mentioned in chapter one. Firstly, atherosclerosis is essentially an intimal disease, in which for early diagnostic tool to be effective, change in the chemical structure within the intima is the primary indicator. Secondary, without puncturing the artery wall, future fiber optical probes will be required to employ Raman

spectroscopy to determine the disease progression by examining mainly the arterial intima. For these reasons, the purpose of this section is to determine whether chemical composition distribution variation within the intima region can be detected by 2D Raman correlation analysis. By employing 2D correlation analysis, if the intima is uniform in composition (only in the case of a perfectly health artery), the correlation plots (synchronous and asynchronous) show no intensity.

6.3.1. 0% (Healthy) Occluded Coronary Artery Intima

The area indicated in figure 6.9 shows a total of 14 sites (25μm apart) where Raman signals were collected within this healthy coronary intima. An almost completely blank (no signal) synchronous plot (figure 6.10(a)) was obtained after employing 2D correlation to the set of spectra. There are minor noise-like peaks appearing in the positive region of the synchronous plots though. They might be due to random shot noise carried from the original Raman spectra and the variation in peak intensity because of unstable laser power. The lack of negative synchronous peak is a good indication that there is no compositional change observed within the selected intima area. The asynchronous plot (figure 6.10 (b)) is also very clean. Hence, the lack of major peaks in both the synchronous and asynchronous plots indicates that the chemical composition of each location is roughly identical.

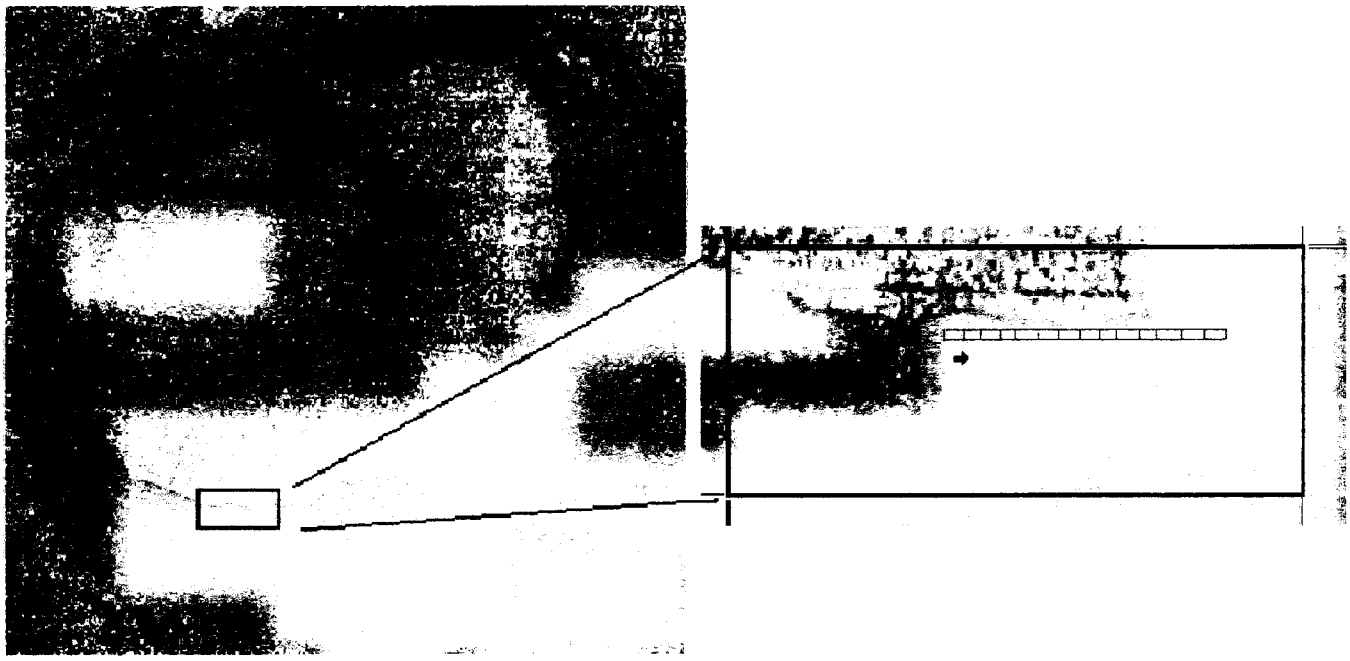
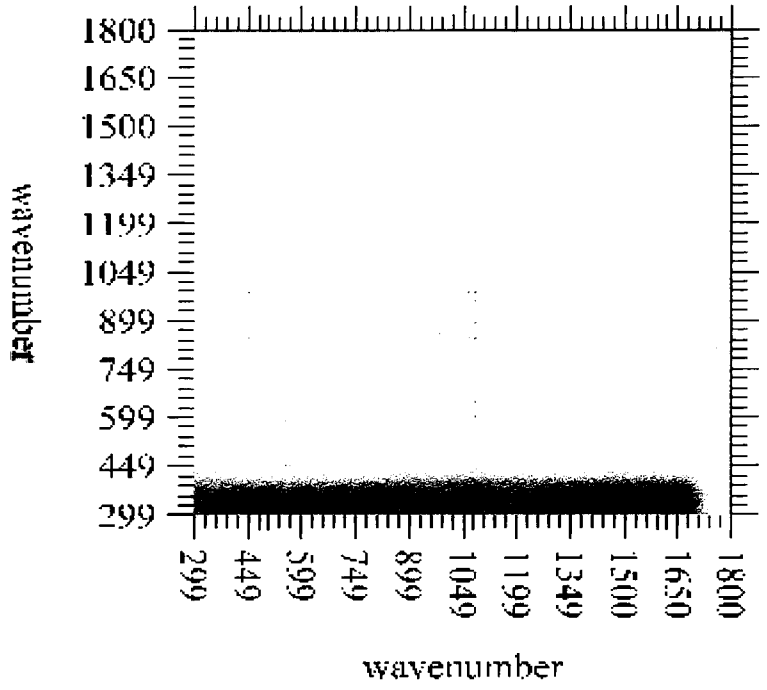


Figure 6.9 Photo of a transverse section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 0%. Spectra were collected from the 14 locations at the center of each grid in 35μm-step (with the direction indicates by the arrow).

a)



b)

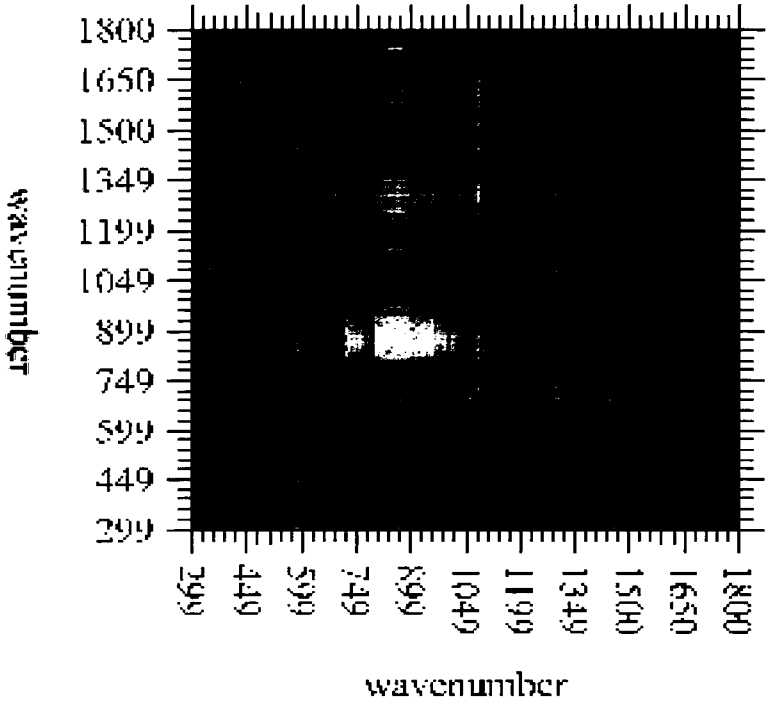


Figure 6.10(a) A 2D contour plot of the synchronous 2D Raman correlation spectrum of the mapping of the 20μm section of the 0% narrowed coronary artery in the spectral region from 300-1800 cm^{-1}. Yellow areas represent the positive peaks, while blue areas represent the negative peaks. (b) A 2D contour plot of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. All peaks are positively shown in blue.

6.3.2. 50-75% Occluded Coronary Artery Intima

The area indicated in figure 6.11 shows a total of nine sites (25μm apart) where Raman signals were collected within arterial intima/plaque of this 50-75% narrowed coronary. In contrast, to the previous case, not completely blank (no signal) synchronous and asynchronous plots (figure 6.12) were obtained after employing 2D correlation to the set of spectra.

Other than the peaks that seem to be spurious, there are some obvious peaks that could result from chemical variations from the different locations of the intima. A chemical composition analysis was performed to this same area and the results did reveal a uniform chemical composition with very little chemical variations in the nine locations (see figure 7.7). Hence, the result supports that 2D correlation can be useful in detection of minor change in chemical composition with the artery.

It is likely that the area selected was a lipid core of a plaque and thus the composition was almost uniform (consisted of mainly triglycerides and cholesterol). If Raman spectra are collected further along concentrically, further composition changes may be observed.

The Raman signatures below 700cm^{-1} resulted in many correlation and discorrelation intensities, which upon further investigation will become very informative to the studying of the tissue composition.

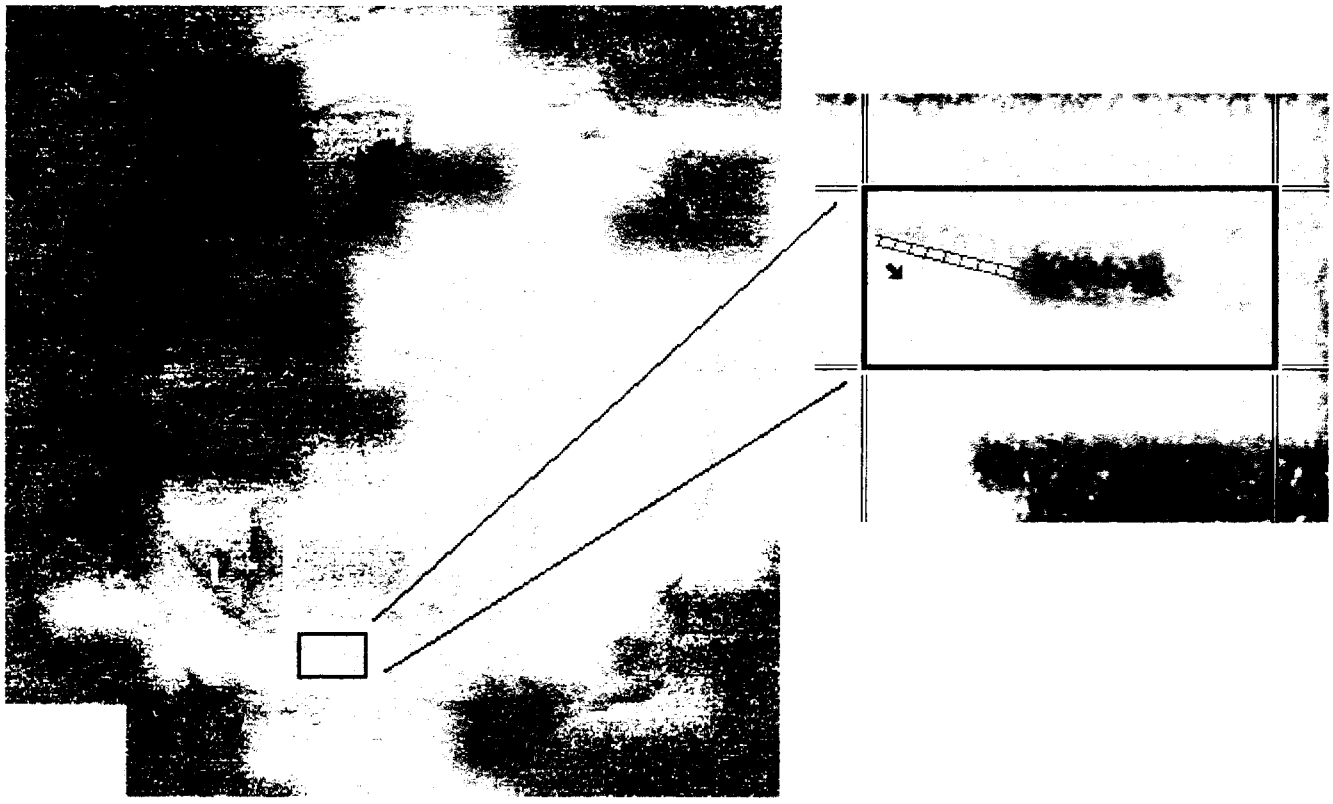
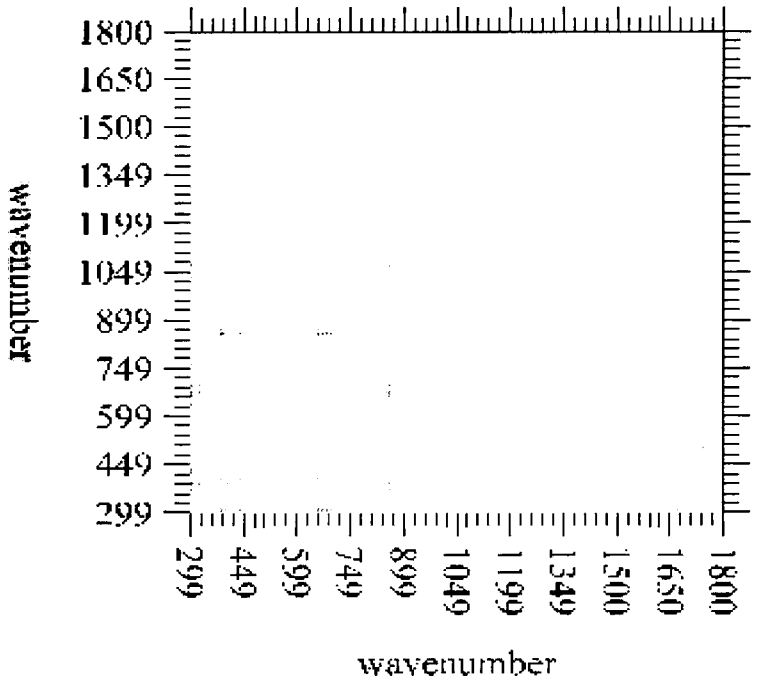


Figure 6.11 Photo of a transverse section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 50-75%. Spectra were collected from the nine locations at the center of each grid in 25μm-step(with the direction indicates by the arrow).

a)



b)

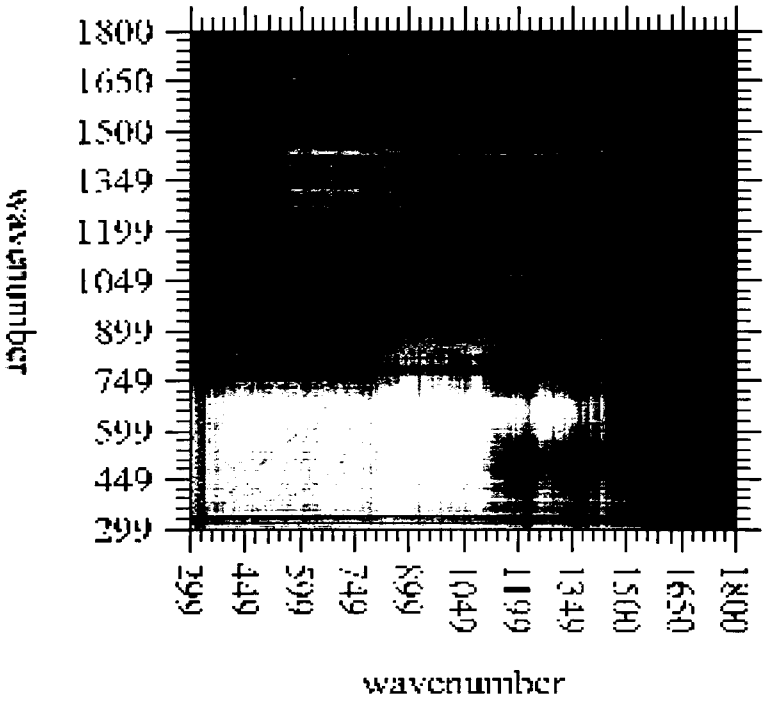


Figure 6.12(a) A 2D contour plot of the synchronous 2D Raman correlation spectrum of the mapping of the 20μm section of the 50-75% narrowed coronary artery in the spectral region from 300-1800 cm^{-1}. Yellow areas represent the positive peaks, while blue areas represent the negative peaks. (b) A 2D contour plot of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. All peaks are positively shown in blue.

6.3.3. 0-25% Narrowed Coronary Artery

The area indicated in figure 6.13 shows a total of 16 sites (25μm apart) where Raman signals were collected within the arterial intima/plaque of this 0-25% narrowed coronary. The synchronous and asynchronous plots (figure 6.14) show a lot of peaks, very different from those in the two previous cases. The plots resulted in here look as complicated as the correlation plots obtained from the cases where along the radius mapping was performed in section 6.2. Thus, the result indicates that the chemical composition concentrically along the intima was not uniform. This may suggest that the area being selected could be an area where endothelial cells damage occurred initially and the formation of plaque was beginning to start. The deposition of different constituents caused the huge variation of the chemical make up within the area. Once again, many correlation and discorrelation intensities resulted below $700cm^{-1}$.

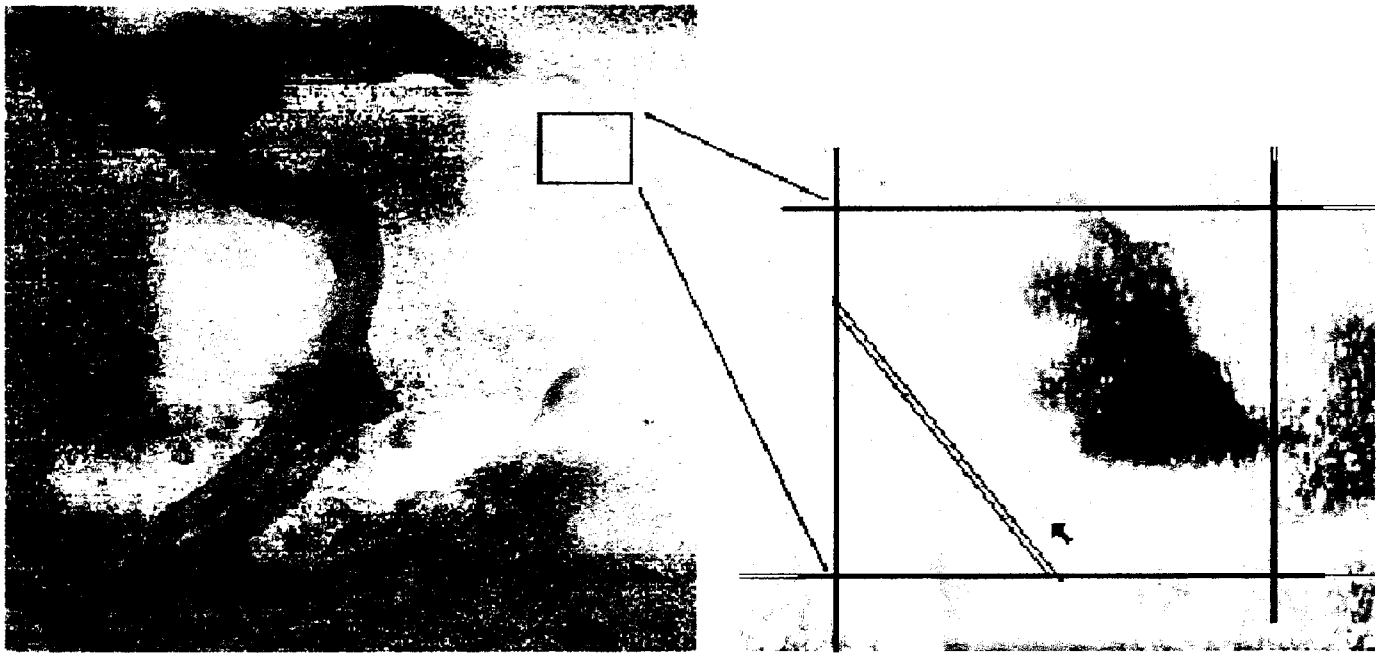
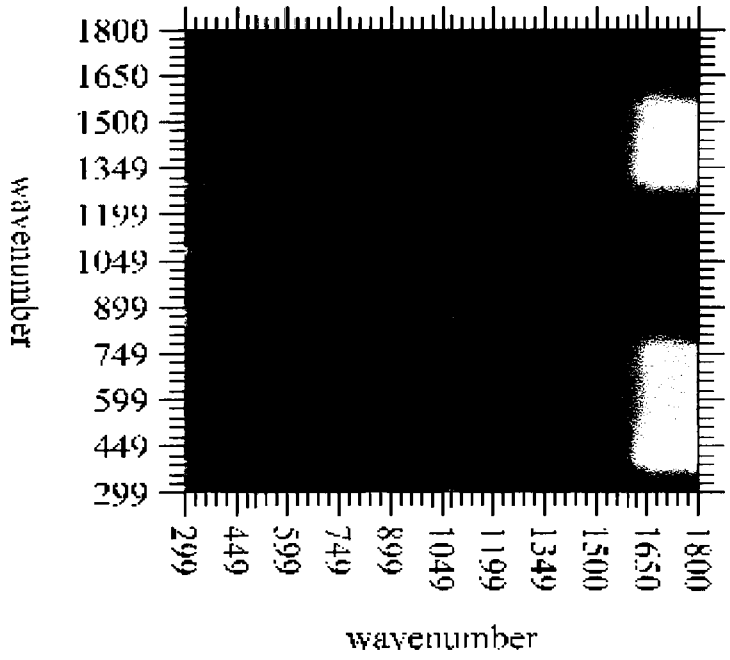


Figure 6.13 Photo of a transverse section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 0-25%. Spectra were collected from the 16 locations at the center of each grid in 25μm-step (with the direction indicates by the arrow).

a)



b)

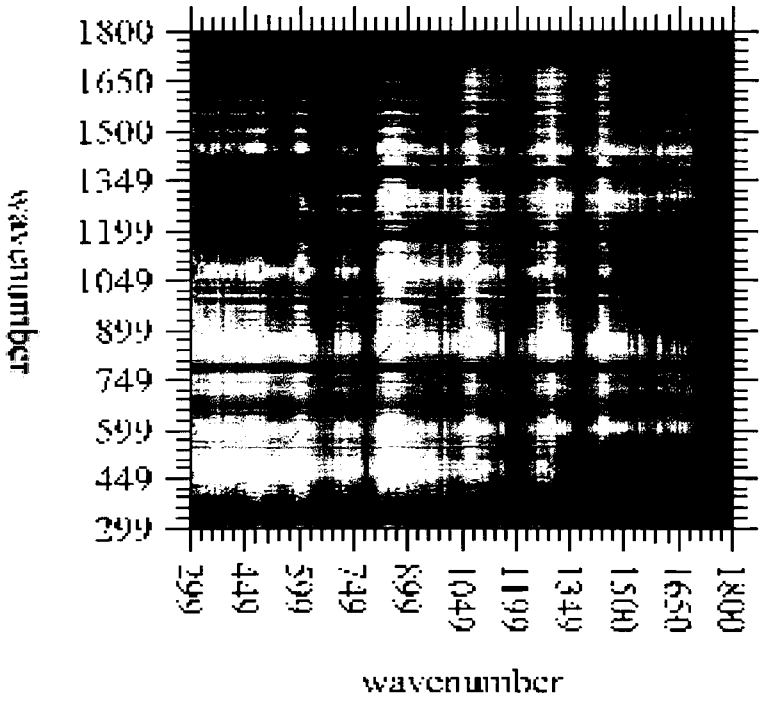


Figure 6.14(a) A 2D contour plot of the synchronous 2D Raman correlation spectrum of the mapping of the 20μm section of the 0-25% narrowed coronary artery in the spectral region from 300-1800 cm^{-1}. Light green areas represent the positive peaks, while dark purple areas represent the negative peaks. (b) A 2D contour plot of the corresponding asynchronous 2D Raman correlation spectrum in the same spectral region. All peaks are shown positively.

6.3.4 Possible Application of 2D Raman Correlation of Coronary Arterial Intima

How will 2D correlations be useful in the diagnosis of atherosclerosis? The power of 2D correlation analysis is its ability to detect subtle change in Raman signatures, and hence minor chemical changes. No other research group has ever employed 2D correlation analysis to classify coronary artery disease. If a library of macroscopic (bigger sampling size than microscopic) Raman spectra can be created from arterial intima using a lager number of atherosclerotic plaques samples of different stages of disease progression, correlation between collected artery spectra from patients and those in the library can reveal the minor difference happens on the compared intima. By determining the correlation and discorrelation information, 2D correlation has the power of diagnosis and assessing the progression and stability of atherosclerosis. Disease can be then classified and early detection can be made possible using solely the 2D correlation information.

The second usage of 2D Raman correlation to the arterial intima, as described earlier, is to complement chemical composition analysis to provide important qualitative information about the order of the appearance of the various constituents.

Chapter 7

Chemical Composition Analysis of Arterial Wall

7.1 Introduction

In order to predict in advance disease progression and the risk of life-threatening complication of atherosclerotic plaques, an understanding of the relationship between a plaque's stability and its chemical composition distribution is desirable [78,79]. Raman spectroscopy offers an attractive tool for surveying the biochemical changes that accompany the development of atherosclerosis. Raman microscopy in automated area-scan mode has already been employed by other research groups to characterize chemical composition distribution in coronary artery, extracting useful information about the constituents present in a tissue site [27,80]. Arterial chemical analysis can provide pathologists quantitative method to extract information to perform a tissue diagnosis by assessing the presence, absence or relative abundance of different constituents.

In this chapter, a general methodology to quantify the chemical composition of the arterial wall is introduced. The measured artery Raman spectrum R(ν) was modeled as a linear combination of individual biochemical Raman line shape $\ell(\nu)$, with fit coefficient x:

$$R(\nu) = \Sigma\, x_i\, \ell_i(\nu) \qquad (7.1)$$

The algorithm employed to calculate the fit coefficients was the singular value decomposition (SVD) method. SVD provides a useful mathematical framework for processing and modeling Raman data. SVD is extremely useful in numerical analysis, for

instance, besides solving the least square problem and data fitting in component analysis introduced in this chapter, SVD has been employed in SNR enhancement [81-83].

7.2 A Short Mathematical Description of Singular Value Decomposition

The singular value decomposition of the observed Raman arterial spectra, an input n x p matrix M (where n is the number of spectra, and p is the number of wavenumbers) can be factorized according to SVD as follow,

$$M = USV^T = CV^T \qquad (7.2)$$

Here U and V are n x n and p x p unitary orthonormal matrices, respectively, and S is an n x p diagonal matrix whose diagonal elements are singular values of the matrix in descending order. Matrix C is called a score matrix, and can be considered a matrix of concentrations of the j chemical components for each spectrum. The use of V^T, the transpose of V (i.e. the matrix obtained by interchanging its row and columns), is required to obey the rules of matrix multiplication [82,84].

SVD of the entire data matrix, M, provides the abstract spectral eigenvectors, U, the diagonal matrix of the singular values, S, and the abstract temporal eigenvectors V, where only the first eighteen columns of U and V and the first eighteen singular values are significant since eighteen pure components were considered in the model. The product

$$A = (S \cdot V^T)^T \qquad (7.3)$$

is a matrix whose elements are the combination coefficient in the equivalent principal component analysis. Any j^{th} column of the M matrix is the linear combination of the spectral eigenvectors in U with the combination coefficients in the j^{th} row of A. A stoichiometric relationship also holds for the rows of A. Since the concentration sum of the intermediates is normalized,

$$A_j \cdot R = 1 \qquad (7.4)$$

where R is an eighteen-component vector.

Thus, once the spectra of the pure components are known, one can construct the calibration matrix A, from which all the concentrations of an unknown sample can readily be obtained.

The program code written to fit the relative contribution of individual constituent to the artery spectra (by SVD) is included in Appendix E.

7.3 Chemical Composition Analysis

Most artery spectra were modeled in the Raman shift range of 560 – 1800 cm^{-1} (although some were modeled in the Raman shift range of 700 – 1800 cm^{-1}), using SVD described in the previous section to generate fractional fit contribution for each of the eighteen pure biochemical components. The model thus provided information about the relative concentration of these chemical components to the Raman spectra of various locations of the artery. The fit contribution of each biochemical component was expressed as a fraction of unity.

Figure 7.1 shows biochemical Raman spectra collected from eighteen commercially available biochemicals used in the model. The laser power and excitation

time used to obtain each spectrum were described in section 4.1 of chapter 4. The spectra were baseline corrected using the cubic spline algorithm.

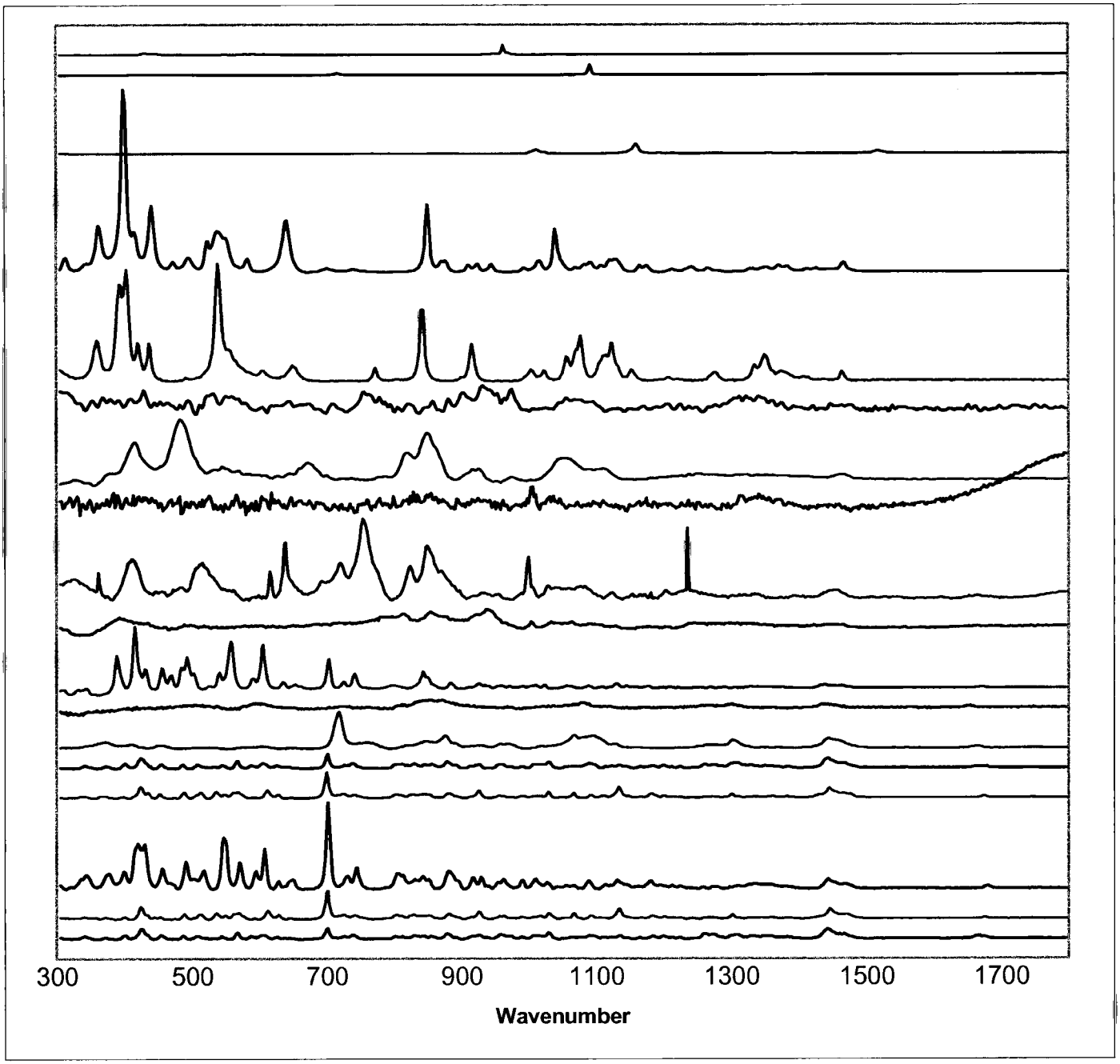


Figure 7.1 Baseline corrected Raman spectra of the eighteen biochemicals used as the basis set. From the bottom to the top these are: cholesterol, cholesteryl oleate, cholesteryl linoleate, cholesteryl palmitate, cholesteryl stearate, L-α-phosphatidylcholine, triolein, 22(S)-hydroxycholesterol, collagen, actin, elastin, myosin, tropomyosin, d-glucose, sucrose, β–carotene calcium carbonate and hydroxyapatite.

7.3.1 Chemical Structures of An Area

The area indicated in figure 7.2 shows a total of 49 sites (7 x 7) where Raman signals were collected within a 100% occluded coronary artery. Each artery spectrum was modeled in the Raman shift range of 700 – 1800 cm^{-1} to generate fractional fit contribution for each of the eighteen chemical components. Figure 7.3 shows the relative amount of each biochemical in each specified locations of the arterial wall.

A trend of appearance/disappearance of chemicals was repeated in the seven rows in which collection was carried out. For example, β-carotene had a higher concentration toward the lumen side and its concentration dropped significantly in the adventitia. On the other hand, there were more triolein (representing triglycerides) in the arterial adventitia compared to the area near lumen. The fitting result of triglycerides was reasonable because adventitia is composed mainly of adventitial fat, which was found to be mainly triglycerides by other studies [45,48].

The dominant components included β-carotene, triglycerides, and hydroapatite. However, β-carotene is a very strong Raman scatterer (large Raman cross section) and does not contain the C-C/ C-H bands near 1440-1455 cm^{-1}, which was employed to normalize all the artery spectra. β-carotene is an intense Raman scatterer that often contributes to coronary artery Raman spectra, but is actually present only in low concentrations (less than 1%) [48]. Hence, its actual concentration should be lower. Additionally, calcium salts (hydroxyapatite and calcium carbonate) as well do not contain the C-C/ C-H bands near 1440-1455 cm^{-1} their actual presence could not be accurately determined. The actual percent of calcium salts composing the arterial wall, as determined by Brands *et al*, varies from 0 to 20% [48]. However, we were able to show

similar outcome that calcium salts could be found in any layers of the artery wall, not exclusively within the plaque/intima area. Proposed methods to avoid the inaccurate fitting results due to the different scattering of individual constituent will be discussed later.

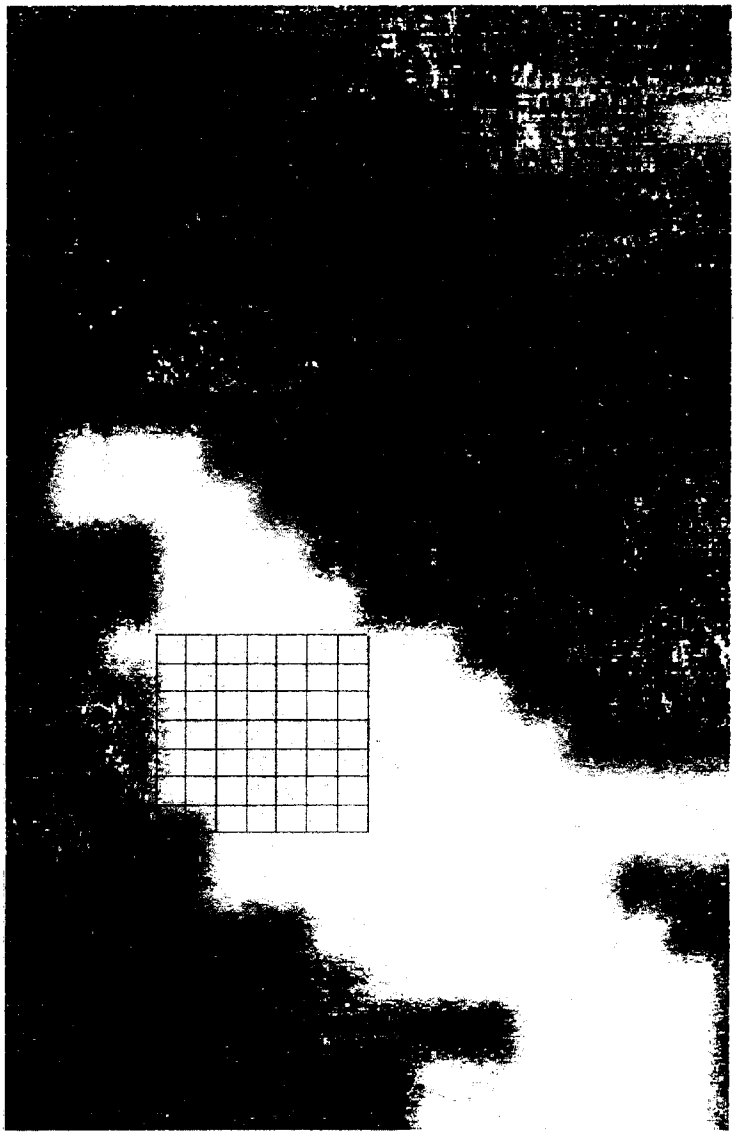


Figure 7.2 A photo of the cross section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 100%. Spectra were collected at the center of each grid from 49 locations, staring from the grid at the top left corner, then row by row unidirectionally from the adventita to the intima Each grid is 100 x 100 μm.

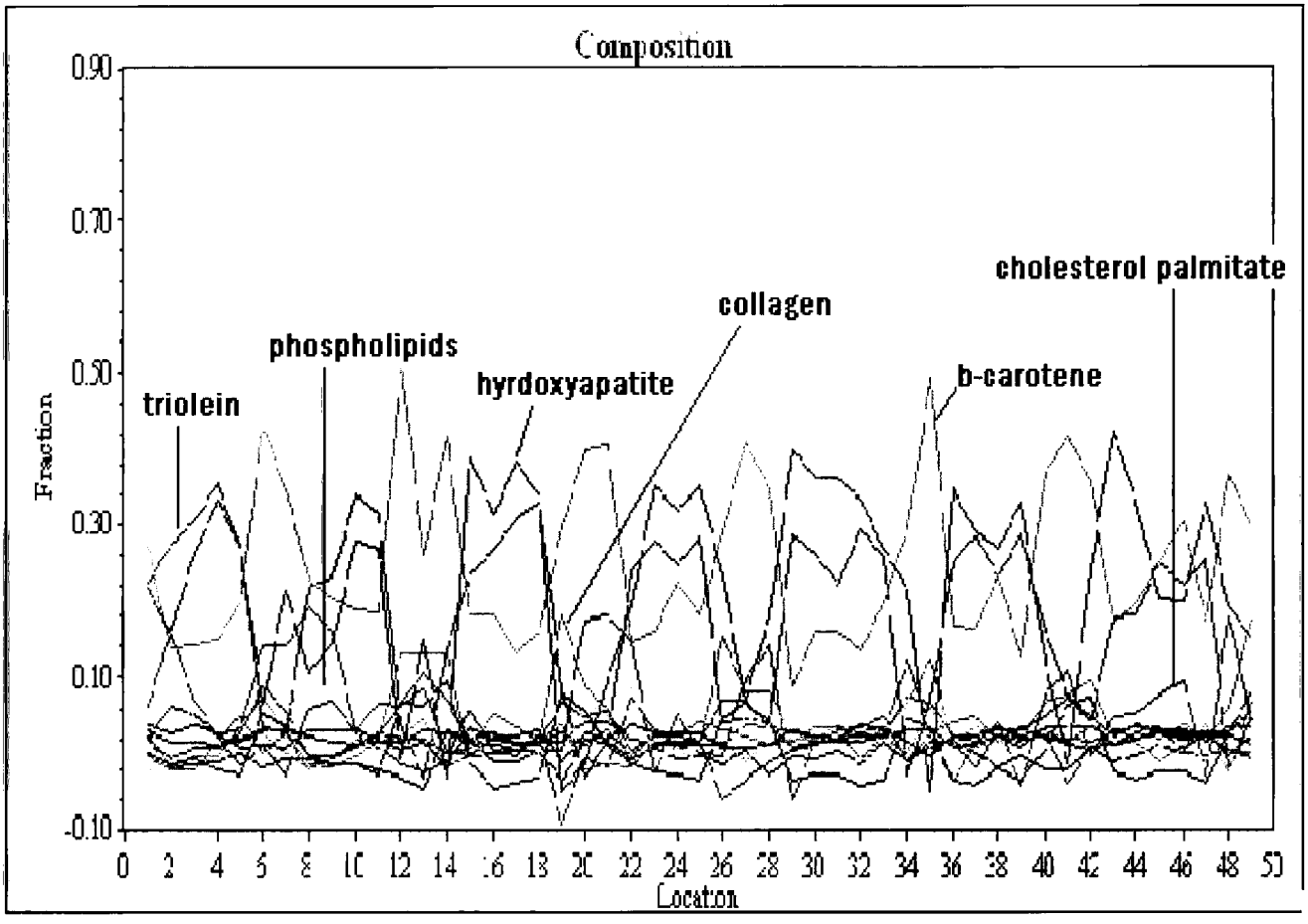


Figure 7.3 The relative fractional amount of the eighteen constituents depicted by the model in each specified location of the arterial wall indicated in figure 7.2.

Besides the three dominating constituents, most other components were determined by the decomposition algorithm to be present in very small amount (less than 5%). However, their presence might have been underestimated. For instance, the total amount of protein determined was not a significant contributor. However, protein contributions were determined by Brands *et al* to be as high as 80% in some locations of an arterial wall [48]. The error could have been due to the low S/R of protein Raman

signals. In addition, some constituents were found to be contributing negatively sometimes. Thus, the accuracy of the algorithm needs to be assessed. Nevertheless, the trend of appearance/disappearance of these chemicals was still able to be shown even though they presence were estimated to be low, for example, cholesteryl palmitate (representing cholesterol ester) and phosphatidylcholine (representing phospholipids) were found more abundant in the area close to the lumen.

Additional error could have been resulted from contributions of biochemical compounds that are not included in the model. For instance, glycosaminoglycans (e.g. hyaluronic acids, chondroitin sulfate, dermatan sulfate, and heparan sulfate, which may contribute 3 % of artery dry mass [47].

7.3.2 Chemical Analysis of the Entire Arterial Radius

A row of spectra was collected (in 100μm steps) over the radius of an artery specimen (100% narrowed), covering the three layers of the coronary artery (intima, media, and adventitia). The area indicated in figure 7.4 shows a total of thirteen sites where Raman signals were collected within this 100% narrowed coronary arterial intima. Each artery spectrum was modeled in the Raman shift range of 700–1800 cm^{-1} to generate fractional fit contribution for each of the eighteen pure chemical components. Figure 7.5 shows the relative amount of each biochemical in each specified location of the arterial wall.

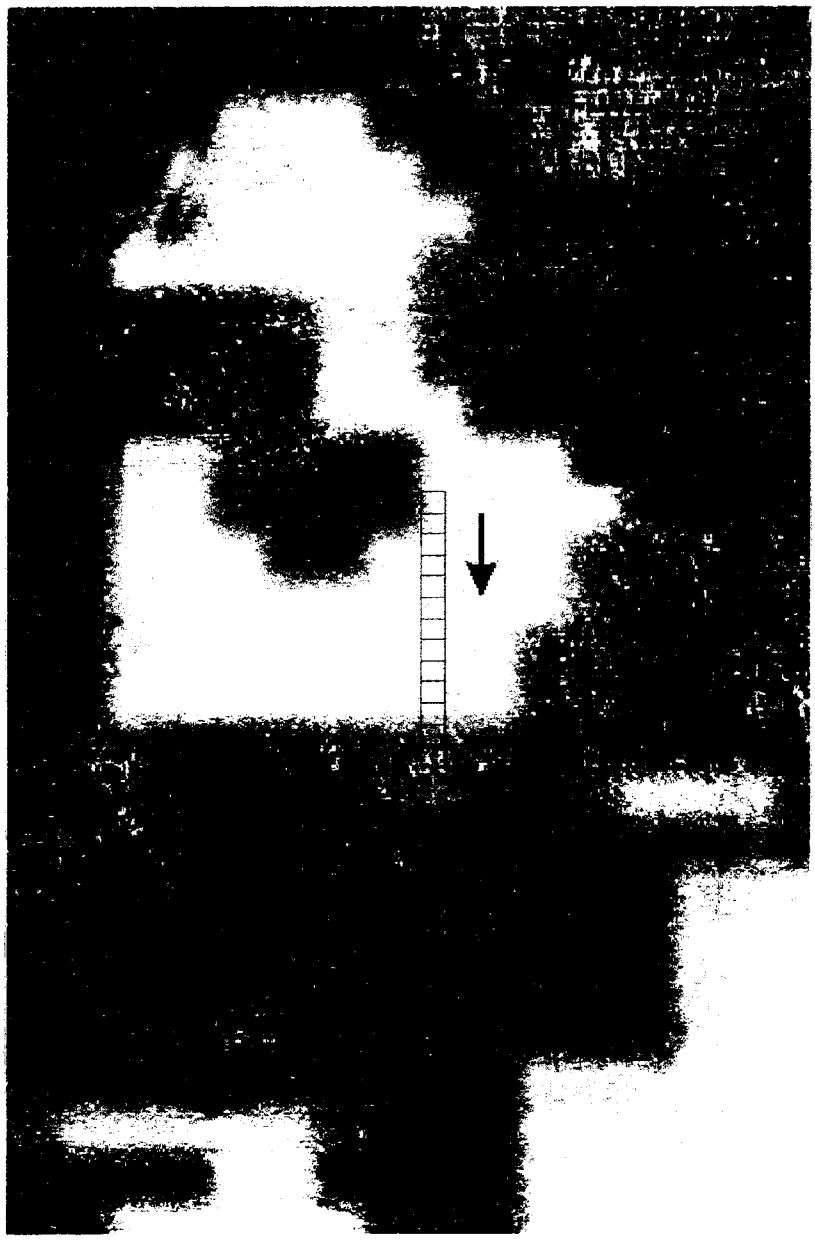


Figure 7.4 A photo of a cross section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 100%. Spectra were collected at the center of each grid from thirteen locations, starting from the adventitia to the intima (direction indicated by the arrow) in 100μm-step.

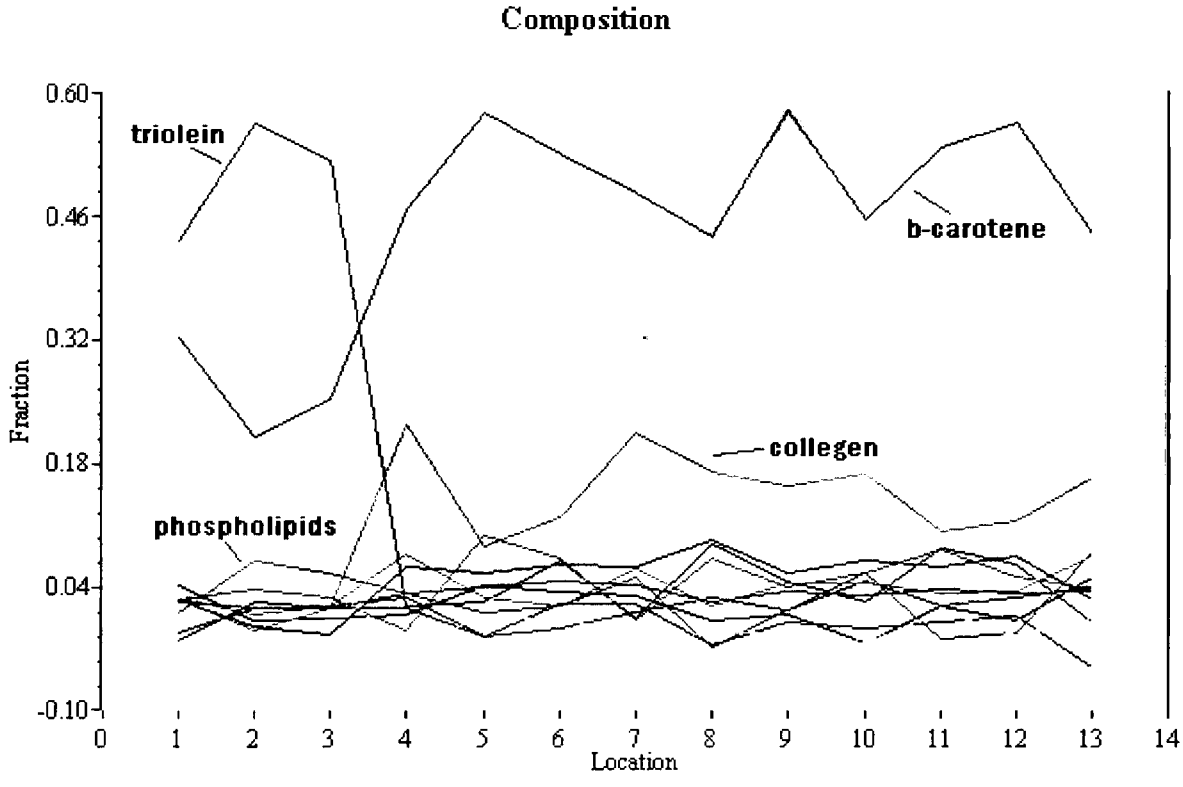


Figure 7.5 The relative fractional amount of the eighteen constituents depicted by the model in each specified location of the arterial wall indicated in figure 7.4.

The spectra were collected from the adventitia to the intima in 100μm steps. As revealed by the plot in figure 7.5, the first three locations consist of a higher concentration of triglycerides than the rest of the area mapped. On the contrary, the amount of collagen and β-carotene were lower in the first three locations, and increased since location number 4. The result was reasonable and could be explained by the fact that location 1 to 3 corresponded to the arterial adventitia, where triglyceride was the dominating

constituent. Location 4 could be assigned the arterial media because collagenous fiber was the major structure making up the medial layer [47]. This artery was assessed as 100% occluded, complete integration of atherosclerotic plaque and the arterial intima occurred. Thus locations 5 to 13 were likely belonging to the plaque/intima, which has high amount of collagen and β-carotene. However, cholesterol and cholesterol esters, which were supposed to be the other major constituents composing a plaque, were unexpectedly low within locations 5 through 13.

7.3.3 Chemical Analysis of the Arterial Intima/Plaque

The area indicated in figure 7.6 shows a total of nine sites (25μm apart) where Raman signals were collected within this 50-75% narrowed coronary arterial intima/plaque. Each artery spectrum was modeled in the Raman shift range of 560–1800 cm^{-1} to generate fractional fit contribution for each of the eighteen chemical components. Figure 7.7 shows the relative amount of each biochemical in each specified location of the arterial wall.

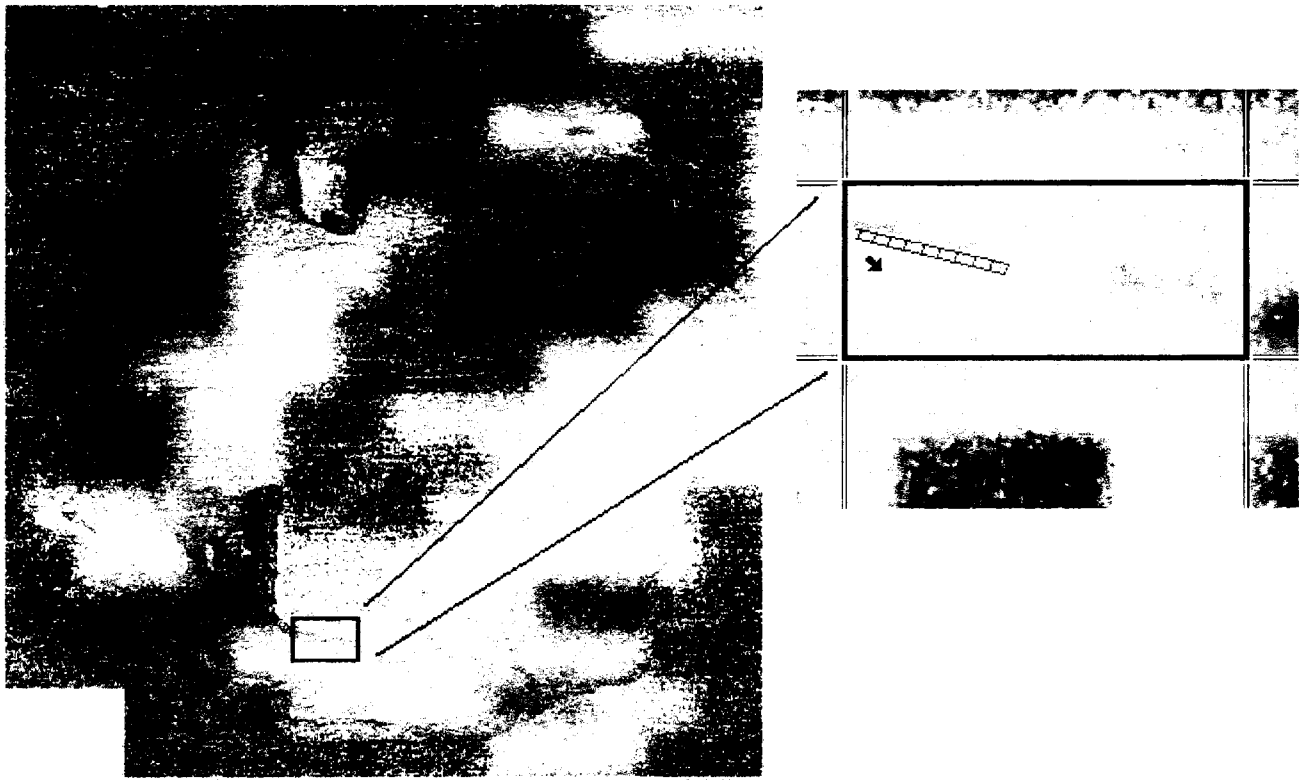


Figure 7.6 A photo of a transverse section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 50-75%. Spectra were collected at the center of each grid (with the direction indicates by the arrow) in 25μm-step.

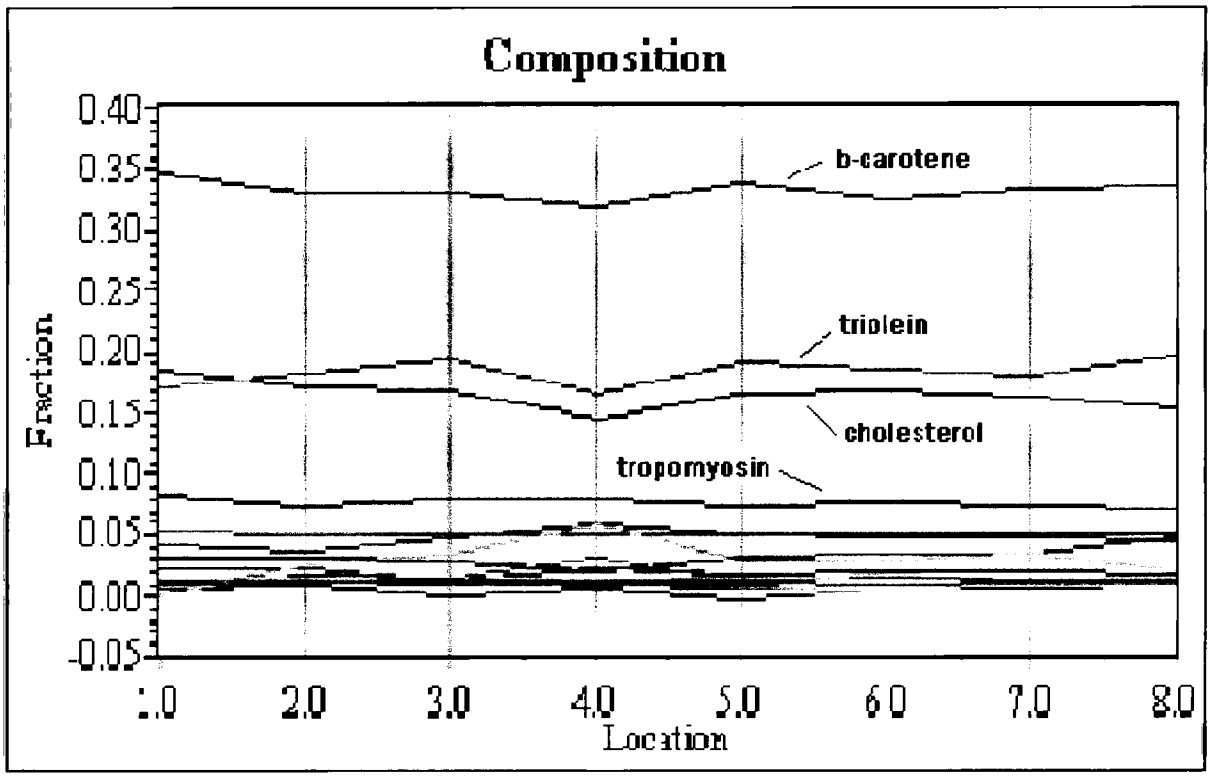


Figure 7.7 The relative fractional amount of the eighteen constituents depicted by the model in each specified location of the arterial wall indicated in figure 7.6.

The decomposition results show a very constant chemical contribution across this small section of the arterial intima. The result here complement the 2D correlation analysis in the previous chapter (section 6.3.2), where no synchronous intensity was observed, indicating no significant chemical change along the area. Free cholesterol and triglycerides (represented by tirolein) were found to be the major components composing this small area of the arterial intima. The presence of cholesterol and triglycerides was

reasonable as these two compounds were two of the major components of a plaque composition [1,2]. Damage to the arterial intima and deposition of fats and cholesterol marks the beginning of the build up of a plaque. As the lesion progresses, cholesterol-rich foam cells accumulate. Thus it makes a lot of sense to find high concentration of free cholesterol and triglycerides at the lumen side of the artery in this 50-75% arrowed artery. Again, the most dominating component was determined to be β-carotene. However, its contribution could have been overestimated.

7.3.4 Chemical Analysis of the Entire Arterial Radius with Fine Steps

In this section, a row of spectra was collected in 20μm steps over the radius of an artery specimen (with 100% occlusion), covering the three layers of the coronary artery (intima, media, and adventitia). Compared to the first two examples, finer step size was performed, and variation within a smaller area of the arterial wall was to be detected. The area indicated in figure 7.8 shows a total of 85 sites where Raman signals were collected within this 100% narrowed coronary arterial wall. Each artery spectrum was modeled in the Raman shift range of 560–1800 cm^{-1} to generate fractional fit contribution for each of the eighteen pure chemical components. Figure 7.9 shows the relative amount of each biochemical in each specified location of the arterial wall.

Figure 7.8. A photo of a cross section of the unstained 25μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 100%. Raman spectra were obtained from 86 locations within the boxed area from the adventitia to the intima in 20μm-step.

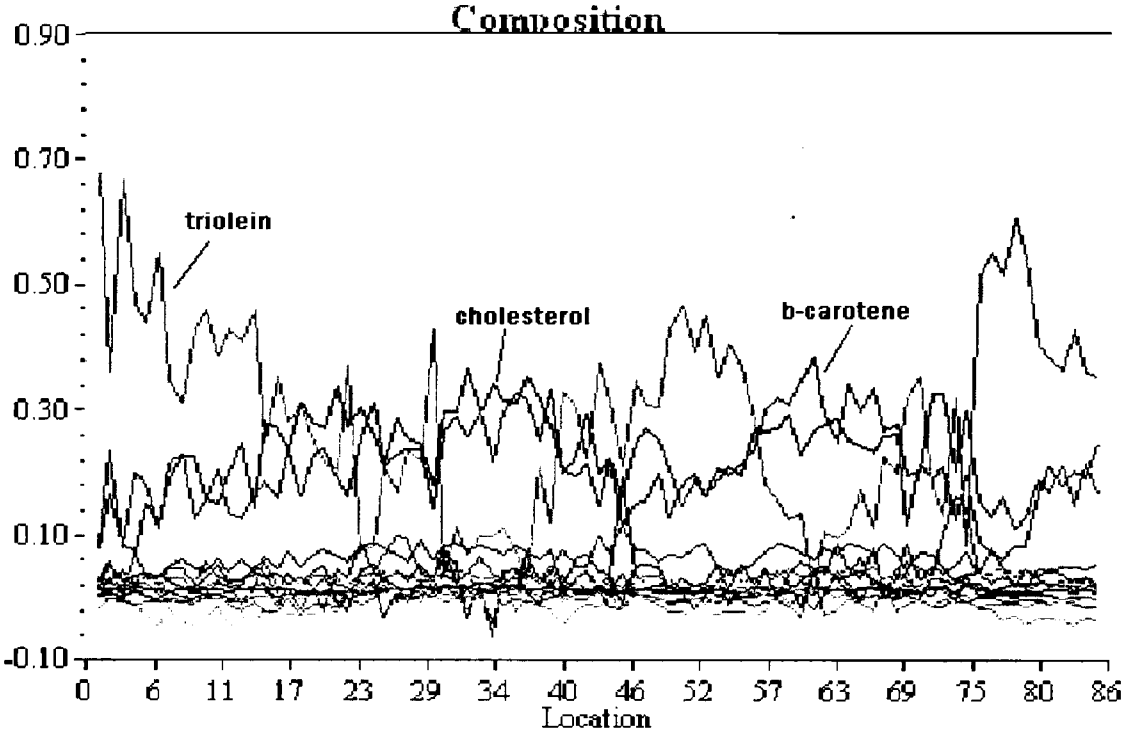


Figure 7.9 The relative fractional amount of the eighteen constituents depicted by the model in each location of the arterial wall indicated in figure 7.8.

The spectra were collected from the adventitia to the intima in 20μm steps. As seen from the plot in figure 7.9, triglyceride concentrations were gradually decreasing during the first sixteen locations (arterial adventitia), and relatively low within the middle portion of the area, and then increased again after location 38 (plaque). On the contrary, the amount of cholesterol and β-carotene were gradually increasing from the adventitia, and they reach their highest concentration at locations 35 to 38, where triglycerides had its lowest concentration. The result was reasonable and could be explained by the fact that location 1 to 30 corresponded to the arterial adventitia, where triglyceride was the dominating constituent. Location 35 to 38 could be assigned the arterial intima/plaque boundary,

where initial deposition of cholesterol and β-carotene occurred. Locations 39 to 85 likely belonged to the plaque, which had high amount of cholesterol and β-carotene. Unfortunately, cholesterol esters, which were expected to be the other major constituents of plaque were not observed. This could have been due to the fact that the total collection time for this complete row was about 5 hours, and melting of the chemicals and their spreading might have happened which may have affected the accuracy of the result.

7.3.5 Chemical Analysis of a 0-25 % Occluded Artery

In this section, a row of spectra was collected in 35μm steps covering the three layers of the 0-25% occluded coronary artery (intima, media, and adventitia). The area indicated in figure 7.10 shows a total of eighteen sites where Raman signals were collected. Each artery spectrum was modeled in the Raman shift range of 560–1800 cm^{-1} to generate fractional fit contribution for each of the eighteen pure chemical components. Figure 7.11 shows the relative amount of each biochemical in each specified location of the arterial wall.

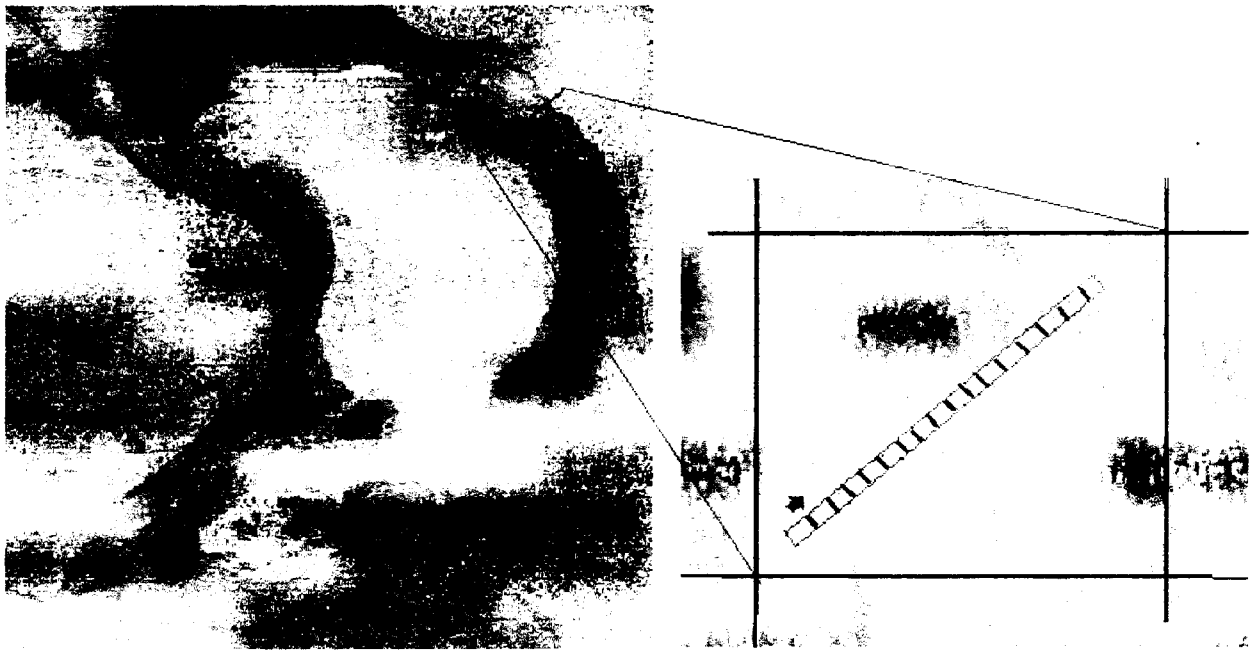


Figure 7.10 A photo of a cross section of the unstained 20μm section of the coronary artery with the amount of cross-sectional area narrowing determined as 0-25%. Spectra were collected from the 17 locations at the center of each grid, starting from the intima to the adventitia (direction indicated by the arrow) in 35μm-step.

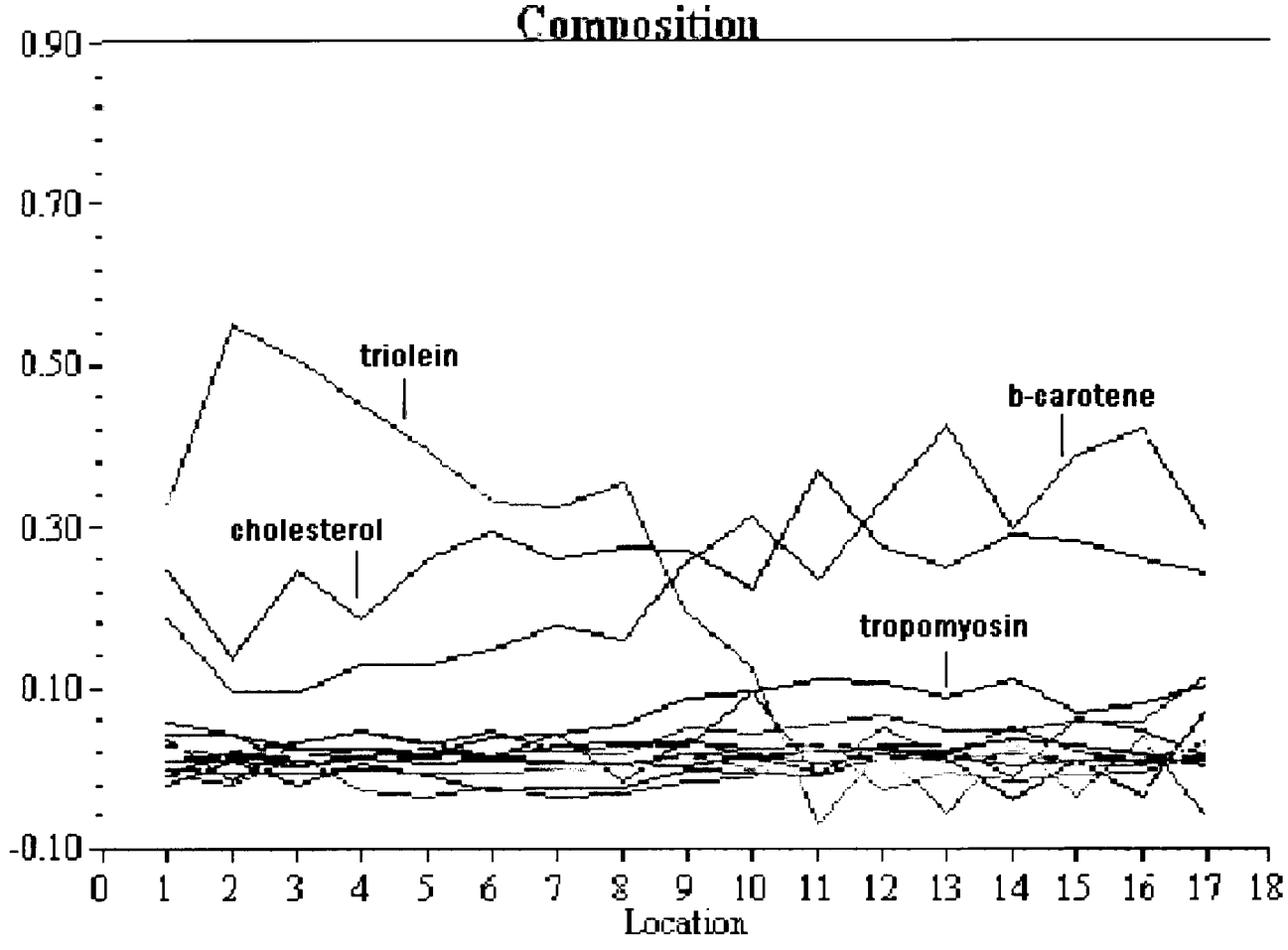


Figure 7.11 The relative fractional amount of the eighteen constituents depicted by the model in each specified location of the arterial wall indicated in figure 7.10.

Locations 1 to 9 were likely belonging to the plaque/intima, which had high amount of triglycerides, cholesterol and β-carotene. The triglyceride concentration dropped significantly after location 9 while on the other hand, cholesterol and β-carotene concentrations increased. Also tropomysin, one of the five proteins included in the decomposition model, increased significantly after location 9. It was reasonable to

observe the protein concentration being higher after location 9 (media and adventita) than in the plaque/intima. Surprisingly, not a huge amount of fat (triglycerides) contribution was found in the arterial adventitia layer as seen in the 50-75 and 100 % occluded arteries discussed earlier. The observed difference might due to the amount of occlusion of the artery. This suggests that the disease progression might involve a change in the composition of arterial adventitia other than plaque/intima.

7.4 Application of the Chemical Composition Results

7.4.1 How Many Constituents to Compose the Basis Set?

The pathological state of human coronary artery can be assessed by quantifying its chemical composition with Raman spectroscopy techniques. The relative weights of the chemical components of the artery are calculated. The question is, how many constituents are required to form the basis set? In addition to the eighteen constituents used in the algorithm, glycosaminoglycans (e.g. hyaluronic acids, chondroitin sulfate, dermatan sulfate, and heparan sulfate), contribute 3% of artery dry mass [47]. In addition, other form of sugars (e.g. fructose) may contribute significantly to the artery composition. Because all these are weak Raman scatterers, they were not included in the model. The more chemical spectra composing the basis set is better, because the more accurate the fitting result that might possibly be obtained. Variations or errors due to contributions of biochemical compounds that are not included in the model can be avoided.

7.4.2 Diagnostic Indicators

The Raman spectra of the many cholesterol esters were very similar. Thus in order to determine the total amount of cholesterol esters at a specific location of the artery, I

believe any one of the cholesterol spectrum alone can represent all the different cholesterol esters very well. The cholesterol linoleate (C18:2) spectrum alone was employed in their many studies to represent the total cholesterol esters [39,47] However, more information can be revealed if more chemicals make up the basis set, for example, the reason I included different cholesterol esters was because the ratio of saturated /unsaturated cholesterol esters is an important indication of the atherosclerosis development. With increasing atherosclerosis development, there is a significant decrease in saturated (palmitic C16:0 and stearic C18:0) esters, and an increase in linoleate (C18:2)/oleate (C18:1) ratio [39].

As seen from previous decomposition results, β-carotene showed an exceptionally high fraction in the artery wall. Because it is a strong Raman scatterer, the contribution of β-carotene might have been overestimated. For the same reason, its fraction fit-contribution was not considered valid also in the previous Raman studies done by other groups when they did their artery composition analysis in their models [47]. However, even β-carotene does not represent a significant weight fraction in arterial wall, its spectrum was used in modeling the artery spectra because carotenoids have been used by others to detect atherosclerotic plaque. Carotenoids were found to be present in greater amounts in plaques than in nonatherosclerotic tissue. In addition, the exact amount of calcium salts (hydroxyapatite and calcium carbonate) was not determined from the model, because they lacked the C-C/ C-H bands in the 1440-1455 cm^{-1} region, which were used to normalize all of the artery spectra. Nonetheless, the presence of calcium salts is another important diagnostic indicator, since its relative weight was found to increase with severity in atherosclerotic plaque [46].

Lipid amount and composition have been shown to vary greatly in atherosclerotic plaques and roughly depends on the grade of the lesion according to its histological classification. The lipid component is linked directly to plaque vulnerability. Total cholesterol, which was found to increase with severity in atherosclerotic plaque, is another useful parameter in classifying the arteries. For instance, the ratio of total cholesterol/unesterified cholesterol can be an important indicator of atherosclerosis development. Lipid plaque contains cholesteryl esters, mainly in a liquid state and a solid, crystalline core, mainly cholesterol. Stabilization of the plaque seems to involve a reduction of the lipid state resulting in an increase of the crystalline state and a hardening of the plaque [41].

7.4.3 Problems with Our Modeling Method

It has been shown that the contribution of proteins in some locations of an artery wall could be as high as 80% by concentration [48]. Using our eighteen-component model, the contributions of the five proteins were not found to be very significant. Failures in the scanned area could be due to a low signal-to-noise ratio of proteins Raman scattering abilities.

The algorithm was programmed such that significant negative contributions from components were ignored. If negative contributions were obtained by SVD, the decomposition was repeated, ignoring the negatively contributed component. However, the reason that caused a constituent to contribute negatively was not known. Once the negative contributed constituent was ignored, an error was introduced due to the non-contribution of the deleted compound.

The spectra from some coronary components are mixtures of closely related molecules that would be difficult to duplicate with commercially available chemicals. The only approach is to extract these compounds from arterial wall and obtain spectra from them. Variations may also be due to peak shifts that might occur when constituents are located close together in the artery as compared to its pure form used in the basis set.

Different scattering of constituents can be solved in two ways. One method is to record power and acquisition time accurately, and then scale the basis spectra and artery spectra obtained according to these recorded parameters. The second way is the proposed use of CaF_2 for artery specimen mounting, will be introduced in the next chapter. Either way can get rid of the problem encountered in our model from the different scattering characteristics of individual constituents.

Chapter 8

Conclusion and Future Directions

8.1 Overview of Conclusion and Future Directions

To conclude this thesis, it is appropriate to briefly summarize the objectives and the progress made toward attaining them. The specific research objectives of this project as outlined previously in the first chapter were i) to characterize the newly installed Raman microscope and to carry out precise point-by-point mapping of pathological specimens, e.g. coronary arteries, ii) to design and carry out spectroscopic mapping of the transverse sections of coronary artery specimens in order to correlate specific spectral features with histologically verified structural and compositional properties, iii) to analyze these data using established spectroscopic methods to obtain qualitative and quantitative information about the chemical structure and composition of plaques, and iv) to utilize the results of this project to prescribe the development of optical fiber probe-based methods for future *in vivo* clinical use.

Previous chapters of this thesis described the equipment set-up, experimental design and preliminary results obtained. This final chapter briefly summarizes the most important insights and conclusions from chapters two through seven. I also point out where improvements may be made, and where future work is likely to be directed in my opinion.

8.2 Equipment

8.2.1 Excitation Wavelength

All spectroscopic data presented in the study were successfully obtained using a 785nm laser excitation source. However, other excitation wavelengths can be made possible with the tunable laser system equipped in the laboratory. It was reported by Feld *et al* in one of their articles that employing an 830nm excitation could yield far better tissue spectra (less fluorescence and better S/N) than 785nm applied in this project [27]. It could be true, as it can be seen in the study that protein spectra contained a lot of noise and fluorescence was still a major distortion in the unprocessed spectra.

Mapping results obtained with different excitation wavelengths can be compared by simply equipping of a tunable bandpass filter and a set of kinematically mounted notch filters that together will enable rapid switching in excitation wavelength to the system with the minimal amount of realignment.

8.2.2 Microscope Optics

The magnification used to select locations to be mapped on the artery cross sections was a standard 5x objective lens equipped with the microscope, giving the visual image an overall 50x magnification (10x magnification eyepiece). However, it was quite difficult to identify morphological structures within the arterial wall, even when specimens were stained. Instead of the 5x objective, a 20x could have been used to identify morphological structures. Nonetheless, an overall of 200x magnification was found to be too powerful in the limited field of view, and morphological structures within the arterial wall were hardly identified. This will limit the possibility of studying morphological structures, which one research group has shown to be a potential diagnostic method very recently

[20]. This research group that worked on collecting Raman spectra from morphological structures used a magnification objective of 10x (an overall magnification of 100x). Thus, more efforts are required to adjust the phase contrast so that the study of morphological structures of the arterial wall becomes more feasible with our microscope in order to identify arterial features using the 5x objective.

8.2.3 Sampling Volume

The volume of sample, which scattered Raman light was collected from the selected area of specimens, is called the sampling volume. This volume under the confocal setting (with the 50x objective lens) was reported in chapter 2 and has been determined by Renishaw and confirmed by another research group that also equipped with an identical Raman microscope.

Future work can be done to characterize the sampling volume or the 'confocal' volume for our system. Polystyrene beads of less than a micrometer in diameter can be used for this purpose [54]. The actual volume may actually be hard to be determined may only be an estimation as mentioned by Neil [85]. For this reason, a characterization of the sample volume can be one of the tasks to be focused on in the future.

8.2.4 More Stable Power

The power generated by the titanium:sapphire under CW (continuous wave) mode was unstable because the beam diameter within the cavity was large and the beam was interrupted at its edges when hitting the optics. Alternatively, if the laser is operated under modelocked (femto-second or pico-second pulses), the beam diameter will become relatively smaller when the laser is producing high intensity modelocked pulses. A more

stable power can be achieved since the beam will pass through the center of optics uninterrupted, such that the power of the laser can remain relatively more constant, while the same Raman effect from samples can be collected.

8.3 Specimens

8.3.1 Specimen Thickness

The thickness of specimen supplied through collaboration in this research project were 20μm and 25μm. There was no difference at all in result observed for the two thicknesses. Different thickness can be tried in future. For example, the penetration of Raman effect can be determined with thinner sections. It was reported by other research groups that a thickness of as thin as 6μm [47] had been utilized in their spectroscopy studies. It must be noted that there is significant amount of difficulty in sectioning highly calcified arteries to very thin slices.

8.3.2 Variety of Specimens

In order for the study to be of future clinical application, a model must be developed to classify plaques into various categories representing the various stages of atherosclerosis. Thus, a lager variety of atherosclerotic plaque samples of different stages of disease progression (stable and unstable plaques of various amount of narrowing) need to be supplied with the generous help from collaborators in hospital.

8.3.3 Mounting on CaF_2 Windows

Even though the use of a metal (aluminium) surface was able to show reduced fluorescence compared to the use regular glass surface, the use of CaF_2 surface is

recommended. The use of CaF_2 was demonstrated to produce less fluorescence than glass or quartz surface in NIR Raman studies [86]. In addition, the Raman spectrum of CaF_2 contains several standard peaks. One of those peaks can be employed to normalize the spectra instead of the C-H/C-C band within 1440-1455 cm^{-1}. Normalization with respect to one of those standard peaks (e.g. 320cm^{-1}) will allow the problem concerning different Raman scattering abilities of individual biochemical to be solved.

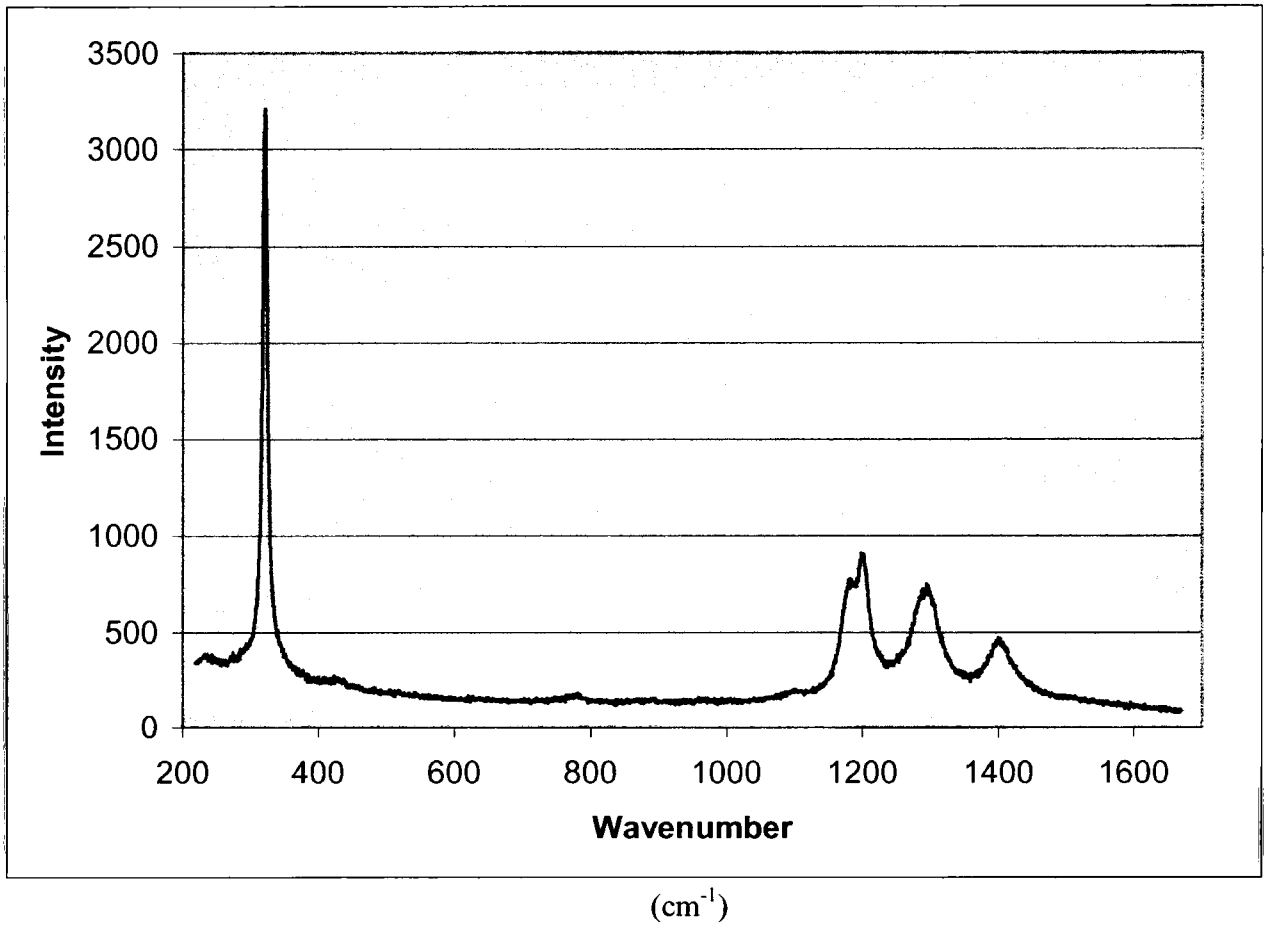


Figure 8.1 Raman spectrum of CaF_2 window with a 785nm excitation laser. Excitation time 30 seconds.

8.3.4 Sample Defrosting

Storage at -80°C was necessary for pathological specimens, however care should have been taken when defrosting the specimens in room temperature. During the time when the specimens were exposed in room temperature, the issue of condensation of the moisture in air caused water to condense as observable water droplets. The accumulation of water might have speeded up dissolving of the biochemical constituents. One easy way to prevent water condensing on the surface of the sample when defrosting is to allow the specimens being gradually brought to room temperature in a desiccator.

8.4 Raman Data Collection

8.4.1 Room Light

As mentioned, larger trials with different thickness and differently occluded stable and unstable atherosclerotic specimens are suggested to be followed. Enabling measurements to be made under normal light conditions in a busy laboratory would facilitate larger trials. The best and easiest way will be by the use of a light-tight stage enclosure box.

8.4.2 Avoid Biochemical Melting

Raman scattered light from each of the specified location on the arterial wall was collected. The excitation time was 30 seconds, which represented the minimum required by the Renishaw Raman system. Together with the grating movement in order that the collection covers the total spectral range from 300-1800 cm^{-1}, a total required time for each location was approximately three and a half minutes. To map the arterial wall with fine steps, the mapping could easily require the specimens be exposed for hours at room temperature. Even though degradation was not observed, melting of biochemicals and the

uncontrollable spreading were unavoidable. To avoid melting, the stage area be maintained cold with a cold cell. Alternatively, the use of a lower resolution (600 or 1200 lines/mm) grating prism is suggested in order to reduce the amount of total collection time.

8.4.3 Sampling Volume

Adjusting the slit width and the size of CCD image area can vary the sampling volume. The most desirable sampling volume is one that matches the size of a single cell, to avoid variations collecting Raman from different cell structures (e.g. cell wall, nucleus etc). The exact imaging volume needs to be determined using the polystyrene bead method as mentioned earlier in section 8.2.3.

8.4.4 Spectral Range

The Raman spectral range has to be selected such that it contains the most useful information regarding the chemical composition of the arterial wall. However, the wider the spectral range, the more the total collection is required, which may affect the accuracy of the result due to biochemical melting or degradation. In addition, the more closer the signals are to the Rayleigh scattered light (at 0 cm^{-1}), the more errors that are introduced, due to errors in baseline correction. On the other hand, if too narrow a spectral range is selected, necessary information may be ignored. Thus, determining of a suitable spectral range to be used in a model that produces accurate diagnostic results is important.

8.5 Raman Spectra Processing

8.5.1 Future Directions in Signal Processing

Many efforts to improve the signal processing method in order to extract the most information possible from a given spectrum become more widespread in spectroscopy. For example, these may include SNR enhancement, noise removal, etc; especially as computing power continues to improve. Thus, there is still a lot of room for improvement in Raman signal processing, including the possibility of real-time signal processing [87]. Furthermore, when an *in vivo* optical-fiber NIR Raman probe is employed in tissue diagnostic applications, background from the fiber optics need to be considered, and subtraction of contribution from the fiber probe will be required to provide quantitatively useful spectra. As well, relating information obtained in the blood-tissue interface to the standard information currently measured in tissue itself will be a possible area that spectra processing might be associated [27].

8.5.2 Use of CaF_2 Windows

Also described earlier in the chapter (section 8.3.3) is the proposed use of CaF_2 slides. Subtraction of the contribution from CaF_2 surface and unwanted background is required as in the case with aluminium metal surface. However, normalization will then be performed before subtraction, with respect to one of its characteristic signals (e.g. $320cm^{-1}$) such that the problem encountered in the chemical composition study caused by different scattering ability of the biochemicals can easily be solved.

8.5.3 Spectrum Smoothing

In 2D correlation analysis, pre-smoothing of the spectra allows the correlation plots to reveal more distinct, more easily identified peaks; however, the process might have thrown away important information especially when Raman signals are generally weak and background is extremely noisy. Thus, there is a trade-off between easier interpretation of 2D correlation studies versus the amount of information lost due to smoothing process, and further decision will be needed to understand this situation.

8.6 2D Correlation Analysis

8.6.1 Application in Diagnosis

There are two important steps in the developing of a model, whether to determine correlations or make diagnoses. The first step, calibration, uses data to determine the values of parameters employed in the model. The second step, validation, applies the model and tests it on new, independent data. If the predictions on the new data set are good, then the model can be considered robust. The power of 2D correlation analysis is its ability to detect minor changes in Raman signatures, and hence slight chemical changes. No other research group has ever employed 2D correlation analysis to classify coronary artery disease. If a library of Raman spectra can be created from arterial intima using a lager number of atherosclerotic plaques samples of different stages of disease progression, correlation between collected artery spectra from patients and those from the library can be used to detect minor change happens on the intima, determine correlation and classify disease and early detection can be a possibility using the 2D correlation information. Of course, large clinical trials are required to confirm these results.

8.6.2 Complement Chemical Composition Analysis

As already mentioned at the end of chapter six, 2D correlation analysis can be performed on small sections of the arterial cross section. Dividing the arterial wall into small subsections and using the information from asynchronous correlation plots, the order of chemical appearance can be determined. The result obtained from 2D correlation can complement to the chemical information obtained from composition analysis. 2D correlation can act as cross validation to the chemical analysis.

8.7 Chemical Composition Analysis

8.7.1 More Accurate Composition Results

Different scattering of constituents can be solved in two ways. One method is to record power and collection time accurately, and then scale the basis spectra and artery spectra obtained according to these recorded parameters. The second way is the proposed use of CaF_2 for specimen mounting. Either way can get rid of the problem encountered from the different scattering characteristics of individual constituent.

8.7.2 Validation of the Decomposition Algorithm

The accuracy and hence the validity of the SVD algorithm used in the study needs to be tested. A known composition mixture of a few constituents can be used to test the validity of the algorithm. In addition, improvements in accuracy are possible if other composition analysis algorithms can be developed and tested individually, such as PCA (principal component analysis), etc.

8.7.3 Model Development

Same two-step model development in the case mentioned in 2D correlation also applies in chemical composition analysis. A calibration step uses data to determine the values of parameters to be employed in the model, and a second validation step applies the model and tests it on new, independent data. If the predictions on the new data set are good, the model can be considered robust [27]. Composition analysis models have been shown by other research groups to allow classifying tissues into three groups, nonatherosclerotic artery, noncalcified plaque, or calcified plaque [46]. Future work can be focused on the development of a model that can classify plaques into stable and unstable. Progress needs to be made in developing modeling techniques for classifying tissue and extracting concentration information.

8.7.4 Morphological Composition Rather Than Chemical Composition

A diagnostic technique that can analyze plaque morphology *in vivo* might prove extremely useful in the clinical evaluation of atherosclerotic lesions. As a complement to chemical analysis, morphological analysis acts also as a quantitative tool to extract information regarding the microscopic constituents present in a tissue sample, and in this case, the morphological structures. Pathologists make a tissue diagnosis by assessing the presence, absence or relative abundance of such constituents or structures (cell type, mineral deposits, etc). Raman microscopy has the potential to provide such morphological information and very interesting area is worthy of further exploration. Similar to the chemical-based analysis, research group were able to classify coronary arteries into nonatherosclerotic artery, noncalcified plaque, or calcified plaque [20].

8.8 Future Application

8.8.1 In Vivo Fiber-Optic Diagnosis

Raman spectroscopy provides the possibility of being a real-time monitor of the chemical makeup of a laser sample. The scattered light from the target would provide a diagnostic indicator of the nature of the diseased site in combination with the performance of *in vivo* optical fiber technology. The ability to provide clinicians with information about the chemical or morphological composition of the tissue in the cubic millimeter at the tip of the probe in a fraction of a second during surgery or a diagnostic procedure is of enormous potential significance, and will essentially bring the insights of pathology into the operating room and laboratory.

In the remote sensing of artery wall, optical-fiber probe is inserted through catheter and the laser is activated at the desired power only when the distal tip of the fiber is directed towards an area of interest, thereby minimizing the probability of perforation of the artery wall. Water and glass exhibits weak Raman scattering while nonpolar bonds (organics) show strong Raman scattering, thus making Raman spectroscopy well suited for fiber optic based analysis of biological samples. Side-looking contact probes that displace blood will be needed to examine arterial walls. Therefore, the distal end of the probe is the most important part of the entire Raman system, and probe's tip design is a very important issue where future development needs to be concentrated on. Various concerns will have to be addressed when designing and characterizing a fiber optic probe, taking into account the turbidity of biological tissue such as developing light collection geometries to improve S/N, eliminating interference from blood, designing the probe-tissue interface, and characterizing depth-ranging information. Otherwise, problems can

arise from the probe itself, causing spectral distortions and limiting light collection and depth resolution [27]. Our laboratory is equipped with modern fiber polishing instruments (ULTRAPOL 1200 series polishing bases and UT-1600 micropositioned polishing heads, Ultra Tec, CA) and the design and development of efficient fiber-optic probes for *in vivo* Raman tissue spectroscopy can be explored. The UBC Raman group has the experience and knowledge in optical fiber probe design, development and characterization. For instance, an efficient angled probe geometry for UVRRS (ultraviolet resonance Raman spectroscopy) was pioneered and patented by the group [84-90]. Finally, the clinical Raman system has to be built to be portable, simple to operate and fully functional in a hospital operating room.

8.8.2 Combined Raman With Other Techniques

The combination of Raman scattering and other methodologies should also be explored. Raman diagnosis combined with other detection techniques have already be proven very useful in the study of atherosclerotic plaques.

In laser angioplasty or angiosurgery, removing the disease site, there is a potential possibility of using Raman scattering as a real-time monitor of the chemical makeup of the target. Raman spectroscopy can serve to provide important feedback during the operation. As the laser removes diseased tissue, the scattered light from the target can provide a diagnostic indicator of the nature of the diseased site. It should be possible to use Raman spectroscopy for feedback control of tissue removed by laser ablation.

It has already been shown possible to combine Raman spectroscopy with existing imaging techniques such as X-ray CT, magnetic resonance imaging (MRI) and intravascular ultrasound (IVUS) [23,25,26]. For example, in the case of atherosclerosis,

combining Raman spectroscopy with IVUS is worthy of exploration. The chemical and morphological information provided by Raman spectroscopy complement the tissue density information provided by these existing imaging technologies. Combined use of Raman and micro-PIXE (particle induced X-ray emission) has been studied by Brands *et al* [48], demonstrated the possibility of providing additional information on the elemental composition of the artery. Another example, where combining autofluorescence microscopy and Raman spectroscopy, the chemical composition of ceroid in atherosclerotic plaque was studied, revealing that heme in blood initiates plaque formation [91].

8.9 Final Concluding Remarks

In summary, a new Raman microscopy system for tissue spectroscopy has been set-up and characterized. Preliminary results for human coronary arteries tissue spectra have been obtained, qualitative (2D correlation) and quantitative (composition analysis) analytical Raman spectroscopy of such specimens has been demonstrated. It is abundantly clear that atherosclerotic plaques diagnosis with portable Raman system is nowhere near complete, and a number of additional steps will be required to realize the development of clinical tools based on Raman technology. However it is likely to be viable, especially given the technical advances occurring apace in the field, and the intense interest shown by many research groups. Raman technology has tremendous clinical potential, and there is substantial room for future advances. My sincere hope is that my work described here continues to be interesting in some ways to others. I anticipate that Raman optical probes will ultimately be used as a practical clinical tool in atherosclerosis diagnosis.

References

[1] Lilly LS. (Ed) (1998) *Pathophysiology of heart disease – a collaborative project of medical students and faculty.* Williams and Wilkins, Baltimore.

[2] Plutzky F., Libby P., Nakamura H. (Eds) (2000) *Second US-Japanese dialogue on lipid disorders and coronary artery disease.* Royal Society of Medicine Press, London.

[3] Jairath N. (1999) *Coronary Heart Disease and Risk Factor Management: a nursing perspective.* W.B. Saunders Company, Philadelphia.

[4] Just H., Hort W., Zeiher A.M. (Eds) (1994) *Artheriosclerosis: new insights into Pathognetic Mechanisms and Prevention.* Springer-Verlag, New York.

[5] International cardiovascular Disease Statistics, American Heart Association, 2002.

[6] Chan YM, Radomski MW, Abdulla AS. Common heart diseases. *WellnessOptions.* 2002; **2(5)**: 18-20.

[7] The Changing Face of Heart Disease and Stroke in Canada, Heart and Stroke Foundation of Canada, 1999.

[8] Heart and Stroke Statistical Update, American Heart Association, 2002.

[9] Cardiovascular Diseases – Prevention and Control. WHO CVD Strategy, 2001/2002.

[10] Kullo IJ, Edwards WD, Schwartz RS. Vulnerable plaques: pathobiology and clinical implications. *Annals of Internal Medicine.* 1998; **129(12)**: 1050-1060.

[11] Cohn P. F. (Ed) (1979) *Diagnosis and Therapy of Coronary Artery Disease.* Little, Brown and Company, Boston.

[12] Schut TCB, Witjes MJH, Sterenborg HJCM, Speelman OC, Roodenburg JLN, Marple ET, Bruining HA, Puppels GJ. *In vivo* detection of dysplastic tissue by Raman spectroscopy. *Analytical Chemistry.* 2000; **72**: 6010-6018.

[13] Salzer R, Steiner G, Mantsch HH, Mansfield J, Lewis EN. Infrared and Raman imaging of biological and biomimetic samples. *Journal of Analytical Chemistry.* 2000; **366**: 712-726.

[14] Woolf N. (1982) *Pathology of Atherosclerosis.* Butterworth Scientific, London.

[15] Gotto A.M. Jr., Robertson A.L. Jr., Epstein S.E., DeBakey M.E., McCollum C.H. (1975) *Atherosclerosis.* A Scope Publication , Kalamazoo.

[16] Ashwell M. (Ed) (1993) *Diet and Heart Disease: a round table of factors, second editon.* Chapman and Hall, London.

[17] van de Poll SWE, Römer TJ, Volger OL, Delsing DJM, Bakker Schut TC, Princen HMG, Havekes LM, Jukema JW, van der Laarse A, Puppels GJ. Raman spectroscopic evaluation of the effects of diet and lipid-lowering therapy on atheosclerotic plaque development in mice. *Arteriosclerosis Thrombosis Vascular Biology.* 2001; **21(10)**:1630-1635.

[18] Conti C. R. (Ed) (1986) *Coronary Artery Spasm: Pathophysiology, Diagnosis, and Treatment.* Marcel Dekker, New York.

[19] Deinum G, Rodriguez D, Römer TJ, Fitzmaurice M, Kramer JR, Feld MS. Histological classification of Raman spectra of human coronary artery atherosclerosis using principal component analysis. *Applied Spectroscopy.* 1999; **53(8)**: 938-942.

[20] Buschman HP, Motz JT, Deinum G, Romer TJ, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, Feld MS. Diagnosis of human coronary atherosclerosis by morphology-based Raman spectroscopy. *Cardiovascular Pathology.* 2001; **10**: 59-68.

[21] Fayad ZA, Fuster V. Clinical Imaging of the high-risk or vulnerable atherosclerotic plaque. *Circulation Research.* 2001; **89**: 305-316.

[22] Buschman HP, Marple ET, Wach ML, Bennett B, Schut TCB, Bruining HA, Bruschke AV, van der Laarse A, Puppels GJ. In vivo determination of the molecular composition of artery wall by intravascular Raman spectroscopy. *Analytical Chemistry.* 2000; **72**: 3771-3775.

[23] Romer TJ, Brennan JF, Puppels GJ, Zwinderman AH, van Duinen SG, van der Laarse A, van der Steen AFW, Born NA, Bruschke AVG. Intravascular ultrasound combined with Raman spectroscopy to localize and quantify cholesterol and calcium salts in atherosclerotic coronary arteries. *Arteriosclerosis Thrombosis Vascular Biology.* 2000; **20(2)**: 478-483.

[24] Kern MJ, Meier B. Evaluation of the culprit plaque and the physiological significance of conronary atherosclerotic narrowings. *Circulation.* 2001; **103(25)**: 3142-3149.

[25] Guo W, Morrisett JD, DeBakey ME, Lawrie GM, Hamilton JA. Quantification in situ of crystalline cholesterol and calcium phosphate hydroxyapatite in human atherosclerotic plaques by solid-state magic angle spinning NMR. *Arteriosclerosis Thrombosis Vascular Biology.* 2000; **20(6)**: 1630-1636.

[26] Peng S, Guo W, Morrisett JD, Johnstone MT, Hamilton JA. Quantification of cholesteryl esters in human and rabbit atherosclerotic plaques by magic angle spinning ^{13}C-NMR. *Arteriosclerosis Thrombosis Vascular Biology.* 2000; **20(12)**: 2682-2688.

[27] Hanlon EB, Manoharan R, Koo T-W, Shafer KE, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasari RR, Feld MS. Prospects for in vivo Raman Spectroscopy. *Phys Med Biol.* 2000; **45**: R1-R59.

[28] Clarke RH, Hanlon EB, Brody H, Isner JM. Detection of calcified deposits in atherosclerotic lesions by laser Raman spectroscopy. *Journal of Raman Spectroscopy.* 1988; **19**: 183-188.

[29] Clarke RH, Hanlon EB, Isner JM, Brody H. Laser Raman spectroscopy of calcified atherosclerotic lesions in cardiovascular tissue. *Applied Optics.* 1987; **26(16)**: 3175-3177.

[30] Klug DD, Singleton DL, Walley VM. Laser Raman spectrum of calcified human aorta. *Laser in Surgery and Medicine.* 1992; **12**: 13-17.

[31] Cooke PM. Chemical microscopy. *Analytical Chemistry.* 2000; **72(12)**: 169R-188R.

[32] Choo-Smith LP, Edwards HGM, Endtz HP, Kros JM, Heule F, Barr H, Robinson Jr JS, Bruining HA, Puppels GJ. Medical applications of Raman spectroscopy: from proof of principle to clinical implementation. *Biopolymers (Biospectroscopy).* 2002; **67**: 1-9.

[33] Mahadevan-Jansen A, Follen M, Ramanujam N, Malpica A, Thomsen S, Utzinger U, Richards-Kortum R. Near-infrared Raman spectroscopy for in vitro detection of cervical precancers. *Photochemistry and Photobioogy.* 1998; **68(1)**: 123-132.

[34] Ferraro J.R., Nakamoto K. (1994) *Introductory Raman Spectroscopy.* Academic Press, San Diego.

[35] Skoog D.A., Holler F.J., Nieman T.A. (1998) *Principles of Instrumental Analysis, fifth edition.* Saunders College Publishing, Philadelphia.

[36] Utzinger U, Heintzelman DS, Mahadevan-Jansen A, Malpica A, Follen M, Richards-Kortum R. Near-infrared Raman spectroscopy for in vivo detection of cervical precancers. *Applied Spectroscopy.* 2001; **55(8)**: 955-959.

[37] Brennan JF, Wang Y, Dasari RR, Feld MS. Near infrared Raman spectrometer systems for human tissue studies. *Applied Spectroscopy.* 1997; **51**: 201-208.

[38] Baraga JJ, Feld MS, Rava RP. In situ optical histochemistry of human artery using near-infrared Fourier transform Raman spectroscopy. *Proc. Natl. Acad. Sci. USA.* 1992; **89**: 3473-3477.

[39] Le Cacheux P, Menard G, Quang HN, Dao NQ, Roach AG, Dron D. Quantitative determination of free and esterified cholesterol concentrations in cholesterol-fed rabbit aorta using near-infrared- Fourier transform-Raman spectroscopy. *Spectrochimica Acta Part A.* 1996; **52**: 1619-1627.

[40] Etz ES, Tomazic BB, Brown WE. Micro-Raman characterization of atherosclerotic and bioprosthetic calcification. *Microbeam Analysis.* 1986; 39-46.

[41] Weinmann P, Jouan M, Dao NQ, Lacroix B, Groiselle C, Bonte J-P, Luc G. Quantitative analysis of cholesterol and cholesteryl esters in human atherosclerotic plaques using near-infrared Raman spectroscopy. *Atherosclerosis.* 1998; **140**:81–88.

[42] Römer TJ, Brennan JF III, Bakker Schut TC, Wolthuis R, Van den Hoogen RCM, Jemeis JJ, Van der Laarse A, Bruschke AVG, Puppels G. Raman spectroscopy for quantifying cholesterol in intact coronary artery wall. *Atherosclerosis.* 1998; **141**:117–124.

[43] Hawi SR, Nithipatikom K, Wohlfeil ER, Adar F, Campbell WB. Raman microspectroscopy of intracellular cholesterol crystals in cultured bovine coronary artery endothelial cells. *Journal of Lipid Research.* 1997; **38**: 1591-1597.

[44] Boustany NN, Manoharan R, Dasari RR, Feld MS. Ultraviolet resonance Raman spectroscopy of bulk and microscopic human colon tissue. *Applied Spectroscopy.* 2000; **54(1)**: 24-30.

[45] Brennan JF, Romer TJ, Lees RS, Tercyak AM, Kramer JR, Feld MS. Determination of human coronary artery composition by Raman spectroscopy. *Circulation* 1997; **96**: 99-105.

[46] Romer TJ, Brennan JF, Fitzmaurice M, Feldstein ML, Deinum G, Myles JL, Kramer JR, Lees RS, Feld MS. Histopathology of human coronary atherosclerosis by quantifying its chemical composition with Raman spectroscopy. *Circulation* 1998; **97**: 878-885.

[47] Buschman HP, Deinum G, Motz JT, Fitzmaurice M, Kramer JR, van der Laarse A, Bruschke AV, Feld MS. Raman microspectroscopy of human coronary atherosclerosis; biochemical assessment of cellular and extracellular morphologic structures in situ. *Cardiovascular Pathology.* 2001; **10**: 69-82.

[48] Brands PJM, van de Poll SWE, Quaedackers JA, Mutsaers PHA, Puppels GJ, van de Laarse A, de Voigt MJA. Combined micro-PIXE and NIR Raman spectroscopic plaque characterisation in a human atherosclerotic aorta sample. *Nuclear Instruments and Methods in Physics Research B.* 2001; **181**: 454-459.

[49] Paddock S. Confocal reflection microscopy: the "other" confocal mode. *BioTechniques.* 2002; **32(2)**: 274-278.

[50] Itoh J. Three-dimensional imaging by confocal laser scanning microscopy (CLSM). *Journal of Advanced Science.* 2001; **13(1 & 2)**: 7-10.

[51] Schrof W, Klingler J, Heckmann W, Horn D. Confocal fluorescence and Raman microscopy in industrial research. *Colloid & Polymer Science.* 1998; **276**: 577-588.

[52] Puppels GJ, Colier W, Olminkhof JHF, Otto C, de Mul FFM, Greve J. Description and performance of a highly sensitive confocal Raman microspectrometer. *Journal of Raman Spectroscopy.* 1991; **22**: 217-225.

[53] Renishaw plc, Spectroscopy Products Division, UK, 2000.

[54] Odziemkowski M, Koziel JA, Irish DE, Pawliszyn J. Sampling and Raman confocal microspectroscopic analysis of airborne particulate matter using poly(dimethylsiloxane) solid-phase microextraction fibers. *Analytical Chemistry.* 2001; **73**: 3131-3139.

[55] Greek LS, Schulze HG, Blades MW, Bree AV, Gorzalka BB, Turner RFB. SNR enhancement and deconvolution of Raman spectra using a two-point entropy regularization method. *Applied Spectroscopy.* 1995; **49(4)**:425-431.

[56] Greek LS. A fiber optic UVRSS instrument for biomolecular investigations. Ph. D. Thesis, The University of British Columbia, 1998.

[57] Rouh A. Delsuc MA, Bertrand G, Lallemand JY. The use of classification in baseline correction of FT NMR spectra. *Journal of Magnetic Resonance , Series A.* 1993; **102**:357-359.

[58] Ruckstuhl AF, Jacobson MP, Field RW, Dodd JA. Baseline subtraction using robust local regression estimation. *Journal of Quantitative Spectroscopy & Radiactive Transfer.* 2001; **68**: 179-193.

[59] Grasselli JG, Bulkin BJ. (Eds) (1991) *Analytical Raman Spectroscopy.* John Wiley & Sons, Inc, New York.

[60] Dietrich W, Rudel CH, Neumann M. Fast and precise automatic baseline correction of one- and two-dimensional NMR spectra. *Journal of Magnetic Resoance.* 1991; **91**: 1-11.

[61] Friedrichs MS. A model-free algorithm for the removal of baseline artifacts. *Journal of Biomoecular NMR*. 1995; **5**: 147-153.

[62] Tanabe J, Miller D, Tregellas J, Freedman R, Meyer FG. Comparison of detrending methods for optimal fMRI preprocessing. *NeuroImage*. 2002; **15**: 902-907.

[63] Lunga GD, Basosi R. A simple method for baseline correction in EPR spectroscopy. *Journal of Magnetic Resonance, Series A*. 1995; **112**: 102-105.

[64] Romanenko SV, Romanenko ES, Kolpakova NA. Use of a spline function of a fractional degree for the description of the base line in the determination of platinum by stripping voltammetry. *Journal of Analytical Chemistry*. 2001; **56(1)**: 51-55.

[65] Zhang Y, Mo J, Xie T, Cai P, Zou X. Application of spline wavelet self-convolution in processing capillary electrophoresis overlapped peaks with noise. *Analytica Chimica Acta*. 2001; **437**: 151-156.

[66] Ren Y, Shimoyama M, Ninomiya T, Matsukawa K, Inoue H, Noda I, Ozaki Y. Two-dimensional Fourier Transform Raman correlation spectroscopy study of composition-induced structural changes in a series of ethylene/vinyl acetate copolymers. *Journal of Physical Chemistry Part B*. 1999; **103**:6475-6483.

[67] Harrington PB, Urbas A, Tandler PJ. Tutorial: two-dimensional correlation analysis. *Chemometrics and Intelligent Laboratory Systems*. 2000; **50**: 149-174.

[68] Geng L, Cox JM, He Y. Dynamic two-dimensional fluorescence correlation spectroscopy. Generalized correlation and experimental factors. *Analyst*. 2001; **126**: 1229-1239.

[69] Jung YM, Czarnik-Matusewicz B, Ozaki Y. Two-dimensional infrared, two-dimensional Raman, and two-dimensional infrared and Raman heterospectral correlation studies of secondary structure of β-lactoglobulin in buffer solutions. *Journal of Physical Chemistry Part B*. 2000; **104**: 7812-7817.

[70] Fu R, Klymachyov AN, Bodenhausen G, Dalal NS. Temperature-jump 2D NMR spectroscopy in crystalline solids: a technique for correlating molecular reorientation across the phase boundaries of an order-disorder lattice. *Journal of Physical Chemistry Part B*. 1998; **102**: 8732-8735.

[71] Pancoska P, Kubelka J, Keiderling TA. Novel use of a static modification of two-dimensional correlation analysis. Part I: Comparison of the secondary structure sensitivity of electronic circular dichroism, FT-IR, and Raman spectra of proteins. *Applied Spectroscopy*. 1999; **53(6)**: 655-665.

[72] Noda I, Dowrey AE, Marcott C, Story GM. Generalized two-dimensional correlation spectroscopy. *Applied Spectroscopy.* 2000; **54(7)**: 237A-248A.

[73] Noda I. Determination of two-dimensional correlation spectra using the Hibert transform. *Applied Spectroscopy.* 2000; **54(7)**: 994-999.

[74] Noda I, Marcott C. Two-dimensional Raman (2D Raman) correlation spectroscopy study of non-oxidative photodegradation of β-carotene. *Journal of Physical Chemistry Part A.* 2002; **106**: 3371-3376.

[75] Ren Y, Murakami T, Nishioka T, Nakashima K, Noda I, Ozaki Y. Two-dimensional Fourier Transform Raman correlation spectroscopy studies of polymer blends: conformational changes and specific interactions in blends of atactic polystyrene and poly(2,6-dimethyl-1,4-phenylene ether. *Macromolecules.* 1999; **32**:6307-6318.

[76] Czarnecki MA. Two-dimensional correlation spectroscopy: effect of band position, width, and intensity changes on correlation intensities. *Applied Spectroscopy.* 2000; **54(7)**: 986-993.

[77] Schulze HG, Greek LS, Barbosa CJ, Blades MW, Turner RFB. Signal detection for data sets with a signal-to-noise ratio of 1 or less with the use of a moving product filter. *Applied Spectroscopy.* 1998; **52(4)**: 621-625.

[78] Tousoulis D, Davies G, Crake T, Lefroy DC, Rosen S, Maseri A. Acute Ischemic heart disease. *American Heart Journal.* 1998; **136(3)**: 382-388.

[79] Libby P, Schoenbeck U, Mach F, Selwyn AP, Ganz P. Current concepts in cardiovascular pathology: the role of LDL cholesterol in plaque rupture and stabilization. *The American Journal of Medicine.* 1998; **104(2A)**: 14S-18S.

[80] Hank ZY, Lee SS, Manson J-AE, Hilborn JG. Method to determine chemical composition distribution on matte coating surfaces using Raman microscopy. *Applied Spectroscopy.* 2000; **54(6)**: 783-794.

[81] Schulze HG. An *in-vivo* resonance Raman neuroprobe. Ph. D. Thesis, The University of British Columbia, 1996.

[82] Iwata T, Koshoubu J. New method to eliminate the background noise from line spectrum. *Applied Spectroscopy.* 1994; **48(12)**: 1453-1460.

[83] Iwata T, Koshoubu J. Minimization of noise in spectral data. *Applied Spectroscopy.* 1996; **50(6)**: 747-752.

[84] Kulcsár A, Saltiel J, Zimányi L. Dissecting the photocycle of the bacteriorhodopsin E204Q mutant from kinetic multichannel difference spectra. Extension of the method of the singular value decomposition with self-modeling to five components. *Journal of American Chmeistry Society.* 2001; **123**: 3332-3340.

[85] Neil JE.Modeling and measuring the effect of refraction on the depth resolution of confocal Raman microscopy. *Applied Spectroscopy.* 2000; **54(6)**: 773-782.

[86] Schuster KC, Reese I, Urlaub E, Gapes JR, Lendl B. Multidimensional information on the chemical composition of single bacterial cells by confocal Raman microspectroscopy. *Analytical Chemistry.* 2000; **72**: 5529-5534.

[87] Schut TCB, Wolthuis R, Caspers PJ, Puppels GJ. Real-time tissue characterization on the basis of *in vivo* Raman spectra. *Journal of Raman Spectroscopy.* 2002; **33**: 580-585.

[88] Greek LS, Schulze HG, Haynes CA, Blades MW, Turner RFB. Rational design of fiber-optic probes for visible and pulsed-ultraviolet resonance Raman spectroscopy. *Applied Optics.* 1996; **35(21)**: 4086-4095.

[89] Greek LS, Schulze HG, Blades MW, Haynes CA.Fiber-optic probes with improved excitation and collection efficiency for deep-UV Raman and resonance Raman spectroscopy. *Applied Optics.* 1998; **37(1)**: 170-180.

[90] Schulze HG, Greek LS, Barbosa CJ, Blades MW, Bree AV, Gorzalka BB, Turner RFB. The measurement of some small-molecule and peptide neurotransmitters with a fiber-optic probe using pulsed-ultraviolet resonance Raman spectroscopy. *Journal of Neuroscience Methods.* 1999; **92**: 15-24.

[91] van de Poll SWE, Schut TCB, van der Laarse A, Puppels GJ. *In situ* investigation of the chemical composition of ceroid in human atherosclerosis by Raman spectroscopy. *Journal of Raman Spectroscopy.* 2002; **33**: 544-551.

```
// Script also for baseline removal and normalization, 5th order polynomial
function splineCorrect(Temp_Spectrum, percentage, Spec_Domain)
assume local;
temp_Spec_Domain = Spec_Domain;
project minFWHM;

flag=1;
temp = Temp_Spectrum;
add = int(percentage*temp.size + 0.5*percentage*temp.size*random()) + minFWHM;
while add <= 0.0 do
   add = int(percentage*temp.size +  0.5*percentage*temp.size*random());
end while;
temp[temp.size+1:temp.size+add] = Temp_Spectrum[temp.size-add+1:temp.size];
x[flag] =1;
y[flag]=1.0;
for j=1 to temp.size - add - 1  step add do
   x[flag]=j;
   y[flag]=temp[j];
   flag=flag+1;
end for;
//for j= temp.size-add to temp.size step 1 do
   x[flag]=temp.size;
   y[flag]=temp[temp.size];
   flag=flag+1;
//end for;
[coefs, intervals] = spline(x, y,  <natcubic>);
Baseline_estimate  = splineEval(Spec_Domain, <natcubic>, coefs, intervals);
//[coefs, intervals] = spline(x, y, <b>, , );
//Baseline_estimate = splineEval(Spec_Domain, <b>, coefs, intervals);
Spec_Domain = temp_Spec_Domain;
return  Baseline_estimate;
end function;



function f_SplineCorrection(rawData, Spec_Domain, start, stop)
assume local;

project FWHM;
project HWHM = int((FWHM+1)/2);
project win_dow = 2*HWHM+1;
project hWHM = 1;
project current_plotHandle;
project Graph_Baseline;
```

```
project Graph_Corrected;
project current_set = 0;
project multiplier = 0.0;
project current_SNR = 0.0;
project Stopping_SNR;
project running_started;
project fraction_of_peak_retained;
project fraction_of_segment_retained;
project stopping_activity;

stopFactor = 0.000000000001;
all_rawData = rawData.columns;
redo_fit = 1;
Data_FWHM = estimate_FWHM(rawData[*,1]);
if Data_FWHM = 0
   then
      Data_FWHM = 2;
   end if;
percentage_init = 2.0*getNumber("Average peak FWHM (channels)for spectrum: ", Data_FWHM)/Spec_Domain.size;
maxVal = 1.0e10;

////////////// CODING ////////////////////////////

        // Fit polynomials to spectrum
while redo_fit = 1 do
   w = {v:1.0};
for i=1 to all_rawData do
temp_Spec_Domain = Spec_Domain;
   current_set = i;
   plotHandle = current_plotHandle[2] + 1;
// f_plot2D(Graph_Baseline, current_plotHandle[2], Spec_Domain, rawData[*,i], "Data", <cartesian>,
//       10, 10, max(Spec_Domain), max(rawData[*,i]), min(Spec_Domain),  min(rawData[*,i]));
   addPlot(Graph_Baseline,Spec_Domain,rawData[*,i]);
   current_plotHandle[2] = plotHandle;
   Tolerance = max(rawData[*,1]);
   limit = 64000;
   multiplier = 3.0;
   count = 0;
   maxCnt = 100*Spec_Domain.size;
   Cnt = 0;
   subtract = 1;
   percentage = percentage_init;
   end_weight = 1.0/10.0;
   start_weight = 1.0/10.1;
```

```
    padding = min(FWHM, (int(0.05*rawData.rows)+1));   // padding weights the fit, anchors spectrum ends
    yPoly = fit(seq(padding, 2*padding, 1), rawData[start:start+padding, i], <poly>, 2, );
    Temp_Spectrum = fitEval(seq(1, padding, 1), <poly>, yPoly);


    //Temp_Spectrum = fill(padding, mean(rawData[start:start+5, i]));
    Temp_Spectrum[padding+1:*] = rawData[start:stop, i];      // region of interest
    Region_of_interest = Temp_Spectrum.size;
    w = fill(Temp_Spectrum.size, 1.0);

    yPoly = fit(seq(1, padding, 1), rawData[rawData.rows-padding+1:rawData.rows, i], <poly>, 2, );
    Temp_Spectrum[Region_of_interest+1:Region_of_interest+padding] = fitEval(seq(padding+1, 2*padding, 1), <poly>, yPoly);

    for j=1 to padding do
       //Temp_Spectrum[Region_of_interest+j] = mean(rawData[rawData.rows-5:rawData.rows,i]);
       w[j]=1.0*(start_weight*(padding+1-j));
       w[Region_of_interest+j]=1.0*(end_weight*j);
    end for;
    Original_Spectrum =Temp_Spectrum;
    Spec_Domain = seq(Temp_Spectrum.size);

    removePlot(Graph_Baseline, );
    current_plotHandle[2] = 0;

oldSNR = 0;
counter = 0;
   while subtract = 1 do
      counter = counter +1;
      if Cnt <= maxCnt then Cnt = Cnt + 1;
         end if;
         // select points for spline
      Baseline_estimate[*,i] = splineCorrect(Temp_Spectrum, percentage, Spec_Domain);
      limit = stdDev(Temp_Spectrum,);
         // and remove outliers above polynomial (i.e. peaks)
      count = 0;
      for j=1 to Temp_Spectrum.size do
         if Temp_Spectrum[j] > Baseline_estimate[j,i]  + multiplier*limit then
            Temp_Spectrum[j] = Baseline_estimate[j,i] + fraction_of_peak_retained*(Temp_Spectrum[j]-Baseline_estimate[j,i]);

            count=count+1;
            end if;
      end for;
      fraction = ((maxVal)/max(Temp_Spectrum - Baseline_estimate[*,i]))^0.5;
      if fraction > 1.0 then fraction = 1.0;
```

```
            end if;
         if fraction  < 1.0 && count == 0 then multiplier = multiplier*fraction;
            percentage = fraction_of_segment_retained*percentage;
            end if;
         maxVal = max(Temp_Spectrum - Baseline_estimate[*,i]);
         limit = stddev(Temp_Spectrum - Baseline_estimate[*,i]);
         current_SNR = maxVal/sqrt(limit);
         if multiplier <= stopping_activity or current_SNR < Stopping_SNR
            then subtract = 0;
            end if;
// ***** New stopping code added here
// ***** Arnel, July 15, 2002
      if counter%500 == 1 then
         if oldSNR == current_SNR then
            subtract = 0;
         else
            oldSNR = current_SNR;
         end if;
      end if;
// ***** New stopping code ends here
            // Plot fitting process
         removePlot(Graph_Baseline, );
         current_plotHandle[2] = 0;
         plotHandle = current_plotHandle[2] + 1;
//       f_plot2D(Graph_Baseline, current_plotHandle[2], Spec_Domain, Temp_Spectrum, "Data", <cartesian>,
//          10, 10, max(Spec_Domain), max(Temp_Spectrum), min(Spec_Domain),  min(Temp_Spectrum));
         addPlot(Graph_Baseline,Spec_Domain,Temp_Spectrum);
         current_plotHandle[2] = plotHandle;
         plotHandle = current_plotHandle[2] + 1;
//       f_plot2D(Graph_Baseline, current_plotHandle[2], Spec_Domain, Baseline_estimate[*,i], "Baseline", <cartesian>,
//          10, 10, max(Spec_Domain), max(Temp_Spectrum), min(Spec_Domain),  min(Temp_Spectrum));
         addPlot(Graph_Baseline,Spec_Domain,Baseline_estimate[*,i]);
         current_plotHandle[2] = plotHandle;
         updateViews();

      end while;
            // Remove baseline first
project Data[start:*,i] = Original_Spectrum[padding+1:Region_of_interest] - Baseline_estimate[padding+1:Region_of_interest,i];
Baseline_estimate[start:*,i] = Baseline_estimate[padding+1:Region_of_interest,i];
project BaselineData = Baseline_estimate;

            // plot corrected spectra
plotHandle = current_plotHandle[3] + 1;
```

```
//f_plot2D(Graph_Corrected, current_plotHandle[3], Spec_Domain, Data[*,i], "Corrected Data", <cartesian>,
// 10, 10, max(Spec_Domain), max(Data[*,i]), min(Spec_Domain),  min(Data[*,i]));
addPlot(Graph_Corrected,Spec_Domain,Data[*,i]);
current_plotHandle[3] = plotHandle;
updateViews(Graph_Corrected);
Spec_Domain = temp_Spec_Domain;
end for;

   redo_fit = getNumber("Repeat with different fitting parameter value (no = 0)?",1);
   if redo_fit = 1
      then
         Stopping_SNR = getNumber("Enter new stopping SNR for residual input data (reduce for a finer fit)", 1.0);
      else if i >= rawData.columns
         then exit while;
      else
         all_rawData = rawData.columns;
      end if;

end while;

running_started = 3;
return Baseline_estimate, Data;
end function;


//////////////  END  ///////////////////////////////////
```

```
function StepSmooth(input)
// This function will take a step function (consisting of 0 and 1s)
// 0 = baseline + signal
// 1 = baseline
// Input: step function as a vector
// Output: step function as a vector
numRows = dim(input);
for j=2 to numRows-1 do
   if input[j]==1 and input[j-1] == 0 and input[j+1]==0 then
      input[j] = 0;
   end if;
   if input[j]==0 and input[j-1]==1 and input[j+1]==1 then
      input[j] = 1;
   end if;
end for;
return input;
end function;
```

```
function PowerSpec(Derivdata)
// A power spectrum is the square of the Fourier transform of a function of time
// Since the spectrum is already a Fourier transform, just need to square each point.
// Input: spectrum as a vector
// Output: power spectrum as a vector
assume local;
numRows = dim(Derivdata);  // Obtain vector size

   for j=1 to numRows do
      PowerSpectrum[j] = Derivdata[j]^2;
   end for;

return PowerSpectrum;
end function;
```

```
// This method applies the baseline removal method #4 to the data.
// 1. Smooth the spectra
// 2. Differentiate the spectra
// 3. Create a power spectrum from the differentiated spectrum
// 4. Create a step funtion identifying the parts of the spectrum that are baseline only
// 5. Smooth this step function out
// 6. Fit a polynomial to the baseline-only parts of the spectra
// 7. Subtract this calculated baseline from the original spectra.

function Method4s(xRaw,yRaw)
assume local;

SmoothedData = MovAvgSmooth(yRaw);
DerivData = DerivVector(SmoothedData);
PowerData = PowerSpec(DerivData);
StepData = threshold(PowerData);
SmoothedStepData = StepSmooth(StepData);
//calcBaseline = Polyfit(SmoothedStepData,yRaw,xRaw);
calcBaseline = Splinefit(SmoothedStepData,yRaw,xRaw,4);
calcSpec = yRaw - calcBaseline;
return calcBaseline,calcSpec;
end function;
```

```
function threshold(input)
// This function takes a power spectrum of a derivative and will decide whether a point is part
// of the baseline or not.  Initial threshold is the mean + 3*standard deviation of the whole data
// set.  Repeat until no points change.
// Input: power spectrum as baseline
// Output: vector containing 1 and 0s. 1 represents part of the baseline, 0 represents a signal+baseline point
assume local;
flag = 0;
numRows = dim(input);
tempinput = input;
while flag == 0 do //repeat until no more values change categories.
   // convert the indices to numbers
   newthresh = mean(tempinput) + 3*stdDev(tempinput);
   temp = 0;
   newinput = 0;
   k = 0;
   for j = 1 to numRows do
      if input[j] < newthresh then
         k = k+1;
         temp[k] = j;  // temp holds the index numbers that are below threshold
      end if;
   end for;
   for m = 1 to k do
      newinput[m] = input[temp[m]];  // put the value of each index stored by temp into newinput
   end for;
   if newinput == tempinput then
      flag =1;
   else
      tempinput = newinput;
   end if;
end while;
output = createVector(numRows,<fill>,0);

for n = 1 to k do
   output[temp[n]] = 1;
end for;
return output, newthresh;  // returns threshold for debugging purposes only, remove after.
end function;
```

```
function Splinefit(stepinput,rawinput,xdata,method)
// This function fits a spline to the data points that are part of the baseline
// Input: vector of 0 (signal + baseline) and 1 (baseline only)
//        vector with raw y values, vector of raw x values
// Output: a vector with the calculated baseline

assume local;
numRows = dim(xdata);
k = 0;
tempx = 0;
tempy = 0;
for j=1 to numRows do
   if stepinput[j]==1 then
      k=k+1;
      tempy[k] = rawinput[j];
      tempx[k] = xdata[j];
   end if;
end for;
smallx = 0;
smally =0;
m = 0;
for i= 1 to dim(tempx)-1 do
   select method from
   case 4:
      if i%(dim(xdata)/8) == 1 then
         m = m+1;
         smallx[m] = tempx[i];
         smally[m] = tempy[i];
      end if;
   case 6:
         if i%(dim(xdata)/8) == 1 then
         m = m+1;
         smallx[m] = tempx[i];
         smally[m] = tempy[i];
      end if;
   end select;
end for;
m = m+1;
smallx[m] = tempx[dim(tempx)];
smally[m] = tempy[dim(tempy)];
[coeffs,intervals] = spline(smallx,smally,<natcubic>);
baseline = splineeval(xdata,<natcubic>,coeffs,intervals);
return baseline;
end function;
```

```
// Generalized 2-D correlation spectroscopy
// Original method by Noda, Appl. Spec. 47: 1329-1336.
// NOTE: Data should contain measured spectra and virtual "mirrored" spectra
// Pancoska et al., Appl. Spec. 53: 655-665.
assume local;

project epsilon = 0.00000001;
project channels = 1;
project Number_spectra = 1;
project filename = "C:\WINDOWS\Desktop\";
project FWHM = getNumber("Full width at half maximum (FWHM)", 8);
project HWHM = int((FWHM+1)/2);
project win_dow = 2*HWHM+1;
project hWHM = 1;                    // for Lorentzian peaks

project acquiredData;
project spec_Domain;
project perturbation_Domain;
project fittedData;
project Corlate;
project Disrlate;
project Corlmap;
project powerSpectra;
project correlationIntensity_profile;

project Graph_Data_sidePanel;
project Graph_Data_topPanel;
project Graph_RegressionCs;
project Graph3D_Fit;
project Graph_preProcessedData;
project Graph_powerSpectrum;
project g_SVD;
project g_PCA;
project Graph_LoadedData;
project _g_IPSF;
project Graph_Control_Discrete;
project Graph_Control_Overlap;
project g_corlIntensity_profileChanges;
project quickMethod;
project synchronous;
project vary_with_perturbation;
project subtract_or_not;
project ratio_or_not;
project run_number;
```

```
project controlPeaks_overlap;
project mirror_data;
project sequence = "initialize";

removePlot(Graph_Data_sidePanel,);
removePlot(Graph_Data_topPanel,);
removePlot(Graph_RegressionCs,);
removePlot(Graph_preProcessedData,);
removePlot(Graph3D_Fit,);
removePlot(Graph_powerSpectrum,);
removePlot(g_SVD,);
removePlot(g_PCA,);
removePlot(g_corlIntensity_profileChanges,);
removePlot(Graph_LoadedData,);
removePlot(_g_IPSF,);
removePlot(Graph_Control_Discrete,);
removePlot(Graph_Control_Overlap,);
updateViews();


// 1 acquire data oe 2 simulate data
acquire_data = -1;
while acquire_data < 0 or acquire_data > 3 do
   quickMethod           = getNumber("Use quick method? (1=yes; 0=no): ", 1);
   synchronous           = getNumber("Emphasize synchronous changes (1 = yes): ",0);
   vary_with_perturbation    = getNumber("Correlate between perturbation endpoints only (=0) or for different perturbation intervals
(=1)? ", 0);
   if vary_with_perturbation    = 1
      then
      project increment  = getNumber("Number of spectra to inlude in each perturbation interval (min = 2)? ", 2) - 1;
      end if;
   mirror_data           = getNumber("Mirror data for curve fitting (yes = 1; no = 0)? ", 0);
   subtract_or_not       = getNumber("Use difference spectra (recommended) and subtract first spectrum from others or subtract
spectrum from the previous interval (no = 0; first = 1; interval = 2)? ", 0);
   ratio_or_not          = getNumber("Use ratios of spectra and ratio all to first spectrum or by interval (no = 0; first = 1;
interval = 2)? ", 0);
   acquire_data              = getNumber("Simulate real spectra = 0; Create spectra systematically (no overlapping peaks) = 1; Create
spectra systematically (overlapping peaks) = 2; Load spectra from file = 3", 1);
   controlPeaks_overlap  = getNumber("Use none (=0), discrete (=1), or overlapping (=2) control peaks? ", 0);

   sequence = "acquire data";
   select acquire_data from
      case 0:
      [spec_Domain, perturbation_Domain, acquiredData] = f_synth_Spectra_Simuln();
```

```
      case 1:
      n_peaks = 3;
      [spec_Domain, perturbation_Domain, acquiredData] = f_synth_Spectra_Discrete(n_peaks);
      case 2:
      n_peaks = 3;
      [spec_Domain, perturbation_Domain, acquiredData] = f_synth_Spectra_Overlap(n_peaks);
      case 3:
      [start, stop, perturbation_Domain, spec_Domain, acquiredData] = f_ImportData();
   end select;

end while;

if vary_with_perturbation = 0
   then
      increment = perturbation_Domain.size - 1;
   end if;
for j = 1 to perturbation_Domain.size-increment do //perturbation_Domain.size-increment
   segment = j;
   // 3 preprocess
   sequence = "preprocess";
   [preProcessed_Perturbation_Domain, preProcessedData] = f_preProcessing(spec_Domain, perturbation_Domain[segment:segment +
increment], acquiredData[*,segment:segment + increment]);
   // 4 correlate
   sequence = "correlate";
   [dataFit, Corlate, Disrlate, Corlmap, Corlate_diagonal] = f_Correlate_all_or_seqtl(spec_Domain, preProcessed_Perturbation_Domain,
preProcessedData, segment);
   fittedData[*,j:*] = dataFit;
   powerSpectra[*,j] = Corlate_diagonal;
   updateViews();
end for;

   // 5 profile
sequence = "profile";
if    vary_with_perturbation    = 1
then  f_profile(spec_Domain, perturbation_Domain, powerSpectra);
end if;

   // 6 archive results
sequence = "archive";
//f_archiveResults();

   // 7 end
sequence = "stop";
```

```
function MovAvgSmooth(xrawdata)
// This function smooths out noise by calculating the moving average
// 2 methods for determining window width:
//   1.  Fixed window width of 8
//   2.  Variable window width using a fraction of the FWHM of the spectra
// input: raw spectral data (one spectrum)
// output: Smooth = smoothed spectrum using a variable width moving average window
//         Smoothfixed = moving average smoothing with a fixed window width 8
assume local;
numRows = dim(xrawdata);  // obtain vector size

// ********* This following segment is the old code that assumes width 8 *******
// for j=9 to numRows-8 do
//    tempData = (vector:
xrawdata[j-8],xrawdata[j-7],xrawdata[j-6],xrawdata[j-5],xrawdata[j-4],xrawdata[j-3],xrawdata[j-2],xrawdata[j-1],xrawdata[j],xrawdata[j+
1],xrawdata[j+2],xrawdata[j+3],xrawdata[j+4],xrawdata[j+5],xrawdata[j+6],xrawdata[j+7],xrawdata[j+8]);
//    Smoothfixed[j] = mean(tempData);
// end for;
// for k=1 to 8 do  // just copy in the first 8 digits to the smoothed data
//    Smoothfixed[k]=xrawdata[k];
// end for;
// for k=numRows-7 to numRows do  // just copy the last 8 digits to smoothed data
//    Smoothfixed[k]=xrawdata[k];
// end for;
// *************** End Fixed Calculation ***************

// The following segment is the code that uses an estimate of the FWHM for width
   windowconstant = 0.4;  // This constant is determines how many FWHMs in the window
   project Data_FWHM;
   project Userinput;
   Avgwindow = toInteger(windowconstant*Userinput);
   if Avgwindow == 1 then  // Minimum window width = 2, otherwise no smoothing occurs
      Avgwindow =2;
   end if;

   tempsize = 2*Avgwindow+1; // size of the vector doing the average.
   tempData = createVector(tempsize,<fill>,0);
   for j = Avgwindow+1 to numRows-Avgwindow do  //do average for each point
      for k = 1 to 2*Avgwindow+1 do  // calculate the window for each point
         tempData[k]= xrawdata[j-Avgwindow-1+k];
      end for;
      Smooth[j]=mean(tempData);  // calculate the average of the window
```

```
    end for;
    for m=1 to Avgwindow do     // copy the fringes used in the moving average calculation
       Smooth[m]=xrawdata[m];
    end for;
    for m=numRows+1-Avgwindow to numRows do
       Smooth[m]=xrawdata[m];
    end for;
// End Variable Window code

return Smooth;//, Smoothfixed;
end function;
```

```
assume local;

project Bases;
project X;
project X_SVD;
project X_PCA;
project X_LU;
project subjectSpectra;
project syntheticSpectra;
project Graph_Composition;
project artificialData;

tolerance = 1.0e-128;
A = Bases;
for n=1 to Bases.columns do
   X.label.row(n)      = Bases.label.column(n);
   A.label.column(n)      = Bases.label.column(n);
end for;
if    artificialData = 1
then  M = syntheticSpectra[*,2:*]+tolerance;//first column is X-values
else  M = subjectSpectra[*,2:*]+tolerance;
end if;

updateViews();

//  temporary  ////////////////
//project M;
//project PC;
//project V;
//project S;
//project A_modified;
//A_modified = A;
//Graph_Composition.axis.x.range.mode = <auto>;
//Graph_Composition.axis.y.range.mode = <auto>;
//////////////////////////


decomposition_method = getNumber("Decomposition method to use: LU (=1); SVD (=2); PCA (=3)",2);
select decomposition_method from

   case 1: //solve using LU decomposition
      A_modified = A;
      X_LU=X[*,1:1]*0.0;
```

```
    for n=1 to Bases.columns do
       X_LU.label.row(n)  = Bases.label.column(n);
       A.label.column(n)      = Bases.label.column(n);
       A_modified.label.column(n) = Bases.label.column(n);
    end for;
    minValue = -1.0;
    while minValue < 0.0 do
       for n = 1 to M.columns do
          X_LU[*, n] = solve(A_modified, M[*,n]);
       end for;

       //eliminate negative contributions
       [A_modified, minValue] = eliminateNegative_spectral_contributions(X_LU, A_modified);
    end while;

case 2: //solve using svd
    A_modified = A;
    X_SVD=X[*,1:1]*0.0;
    for n=1 to Bases.columns do
       X_SVD.label.row(n) = Bases.label.column(n);
       A.label.column(n)      = Bases.label.column(n);
       A_modified.label.column(n) = Bases.label.column(n);
    end for;
    minValue = -1.0;
    while minValue < 0.0 do
       for n = 1 to M.columns do
          [y, residual, rnk] = solve(A_modified, M[*,n], <leastSq>, <svd>, tolerance);
          X_SVD[*, n] = y;
       end for;
       //eliminate negative contributions
       [A_modified, minValue] = eliminateNegative_spectral_contributions(X_SVD, A_modified);
    end while;

case 3: //use PCA and solve for bases in PCA vectors
    A_modified = A;
    X_PCA=X[*,1:1]*0.0;
    for n=1 to Bases.columns do
       X_PCA.label.row(n) = Bases.label.column(n);
       A.label.column(n)      = Bases.label.column(n);
       A_modified.label.column(n) = Bases.label.column(n);
    end for;
    minValue = -1.0;
    [PC, S, V] = SVD(M);
    while minValue < 0.0 do
```

```
            for n = 1 to PC.columns do
               [y, residual, rnk] = solve(A_modified, PC[*,n], <leastSq>, <svd>, tolerance);
               X_PCA[*, n] = y;
            end for;
            //eliminate negative contributions
            [A_modified, minValue] = eliminateNegative_spectral_contributions(X_PCA, A_modified);
         end while;
end select;
```