Development of an Intracoronary Raman Spectroscopy System

by

Alexandra H. Chau

S.B., Massachusetts Institute of Technology (2002)
S.M., Massachusetts Institute of Technology (2004)

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 2009

Author .....
Department of Mechanical Engineering
May 18, 2009

Certified by .....
Guillermo J. Tearney
Associate Professor of Pathology, Harvard Medical School
Thesis Supervisor

Certified by .....
Peter T. So
Professor of Mechanical and Biological Engineering
Thesis Supervisor

Accepted by .....
David E. Hardt
Chairman, Committee on Graduate Students

Development of an Intracoronary Raman Spectroscopy System

by

Alexandra H. Chau

Submitted to the Department of Mechanical Engineering
on May 18, 2009, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Abstract

Atherosclerosis is a leading cause of death in the United States, with 1 in 5 deaths (500,000 annually) attributable to coronary artery disease alone. While the disease processes are not completely understood, it is believed that patient risk depends on a variety of factors including lesion structure, biomechanical behavior, and morphological and chemical composition. Raman spectroscopy, based on spectral analysis of inelastically scattered photons, is a nondestructive technique that yields detailed information about the chemical composition of the sample being interrogated. Intracoronary Raman spectroscopy can be performed via the use of a flexible, small diameter (< 2 mm) optical fiber probe to guide light to and from the arterial wall *in situ*. The fact that Raman scattering has inherently low signal intensity, combined with the need for a small diameter probe, makes it difficult to develop a probe with sufficient signal-to-noise for robust plaque diagnosis. This thesis addresses two approaches for increasing SNR: increasing Raman signal intensity and optimizing probe design.

While most biological applications of Raman spectroscopy have been performed in the "fingerprint" region (Raman shifts between 400 and 1800 cm^{-1}), the high wavenumber region ($2700 - 3100$ cm^{-1}) offers distinct technical advantages, including increased Raman signal relative to the fluorescent background and potentially less fiber background. However, the high wavenumber region may yield different molecular information and thus have different diagnostic capability. In this thesis, we develop a benchtop Raman system capable of acquiring Raman spectra in both wavenumber regions. In contrast to previous work, which focused on plaque characterization based on the Raman spectrum from a single site within the plaque, our system utilizes a line imaging paradigm, in which we acquire Raman spectra at fixed intervals across the full cross-section of the plaque, creating a Raman line image. We use this benchtop system to acquire a database of Raman line images and corresponding histology for over sixty plaque specimens. Using this database, we compare the diagnostic capability of fingerprint and high wavenumber Raman spectroscopy for plaque characterization.

Because of the small size requirement for an intracoronary probe, it is important to optimize the optical probe design to maximize collection efficiency and thus increase SNR. We develop and experimentally validate a simulation technique for modeling Raman probe performance (collection efficiency and sampling volume), as an aid to optimizing probe design. We also fabricate a 1.5 mm diameter probe and demonstrate it *in vivo*, using a human-swine xenograft model, in which diseased human coronary arteries are grafted onto a living swine heart. The results of this thesis provide insight on two approaches toward achieving a clinically viable intracoronary Raman spectroscopy system.

Thesis Supervisor: Guillermo J. Tearney
Title: Associate Professor of Pathology, Harvard Medical School

Thesis Supervisor: Peter T. So
Title: Professor of Mechanical and Biological Engineering

Acknowledgments

The road to completing a PhD is generally long and fraught with challenges, but I am grateful that my road has not been solitary. I am indebted to my advisors Prof. Guillermo J. Tearney and Prof. Brett E. Bouma for providing excellent guidance and for always pushing me to exceed my expectations. I am lucky to have stumbled into their lab at the Wellman Center for Photomedicine as an undergraduate so many years ago. Prof. Tearney and Prof. Bouma have created an environment that is impressive for both the quality of the research and the quality of the researchers. In this nurturing environment, overflowing with both hard work and heart, I feel I have grown from a naive undergraduate to a competent researcher.

The work presented in this thesis forms a subset of a larger Raman program, overseen by Dr. Joseph Gardecki, and in the early years, Dr. Jason Motz, both of whom gave me daily mentorship and friendship for years. I am especially grateful to Joe for always making himself available to help with anything, big or small, and for our long discussions filled with advice on research and advice on life.

I could not have completed this thesis without technical contributions from additional colleagues, and I thank Bill Farinelli, Prof. Seemantini Nadkarni, and Dr. Milen Shishkov, who have provided relatively small (to them) but key (to me) contributions to my thesis research project, in moments of crisis.

Within the lab, I have found that I can learn something from everybody, and I thank Dr. Alberto Bilenca, Dr. Nicusor Iftimia, Priyanka Jillella, Dr. DongKyun Kang, Dr. Linbo Liu, Dr. Hyle Park, Dr. Leilei Peng, Dr. Wang-Yuhl Oh, Dr. Melissa Suter, Prof. Ben Vakoc, and Dr. Hongki Yoo, for enlightening me in technical discussions, for being my friends, and for making me laugh. I also thank my SM thesis mentor, Dr. Raymond Chan, from whom I learned a lot of about research, life, and MATLAB. In addition, I am grateful to my fellow graduate students in the lab, Brian Goldberg and Dr. Caroline Boudoux, for their commiseration on the challenges of being a graduate student off-campus.

I also thank my thesis committee members Prof. Peter T.C. So and Prof. Forbes Dewey for their support and research advice.

I believe that teaching is an important part of the graduate student experience, and I am grateful to Prof. George C. Verghese and Prof. Franz X. Kaertner for giving me

the opportunity to see what classes are like from the other side, as a grader and a teaching assistant. I especially thank Prof. Kaertner for giving me the opportunity to teach tutorials as a teaching assistant for Fundamentals of Photonics, and for his supportive feedback to help improve my teaching skills.

Finally, I thank my family and friends. I am eternally grateful for the love and support of my parents, Peter C. Chau and Kim Q. Chau, who have always put my education first, and who have always strived to instill in me the confidence and the work ethic needed to succeed. In their continual study of their personal hobbies, they provide daily examples of how nothing comes easy without careful study, practice and attention to detail, not even whitewater kayaking or painting. In addition, my brother, Phu H. Chau has motivated me throughout my education. I give thanks to Dr. Jennifer J. Yu, who has been one of my best friends since we started at MIT together as freshmen back in 1998 and with whom I have shared a lot of laughter and (not so many) tears.

My graduate years at MIT have been more enjoyable and more memorable for being shared with my fiancé, Dr. Baris I. Erkmen. He has helped me grow professionally through technical discussions about our research, and personally, with his unflagging confidence in me and his continual encouragement and support. I am grateful to him for all the giggles and the sense of calm that he brings, and I thank him for sharing my life.

This work has been supported by an NIH Kirschstein NRSA Predoctoral Fellowship (1F31EB007169), an NIH Research Project Grant (NIH R01HL093717), and Prescient Medical, Inc.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Over 16 million Americans are afflicted with coronary heart disease (CHD), and approximately one of every five deaths in the United States is caused by CHD, or ~500,000 deaths annually [1]. In the United States, the total economic burden of CHD is ~$165 billion per year in direct and indirect costs. Atherosclerosis is a complex disease, characterized by the development of atheromatous plaques within the arterial wall. When a plaque is disrupted, thrombogenic factors in the atheroma come into contact with the blood, initiating the clotting cascade. The resulting thrombus can occlude flow, cause ischemia, and may ultimately result in myocardial infarction and stroke [2]. The mechanism behind plaque progression and disruption and how it relates to patient risk is not well understood, but lesion structure, molecular and chemical composition, biomechanical behavior, and functional state are believed to be important. There is a growing need for diagnostic technologies that can study these factors, including their clinical significance and temporal evolution, to gain understanding of the disease processes and to assess treatment efficacy.

1.1 Atherosclerosis

The normal human coronary artery is composed of three layers, the intima, media, and adventitia (Figures 1-1 and 1-2). Endothelial cells line the lumenal surface of the intima, where they regulate the passage of molecules between the blood and the arterial wall. The intima is a thin layer of connective tissue, primarily collagen and elastin, with a small number of smooth muscle cells. A layer of elastic tissue, the internal elastic lamina, serves as a barrier between the intima and the media. The media contains tightly packed layers

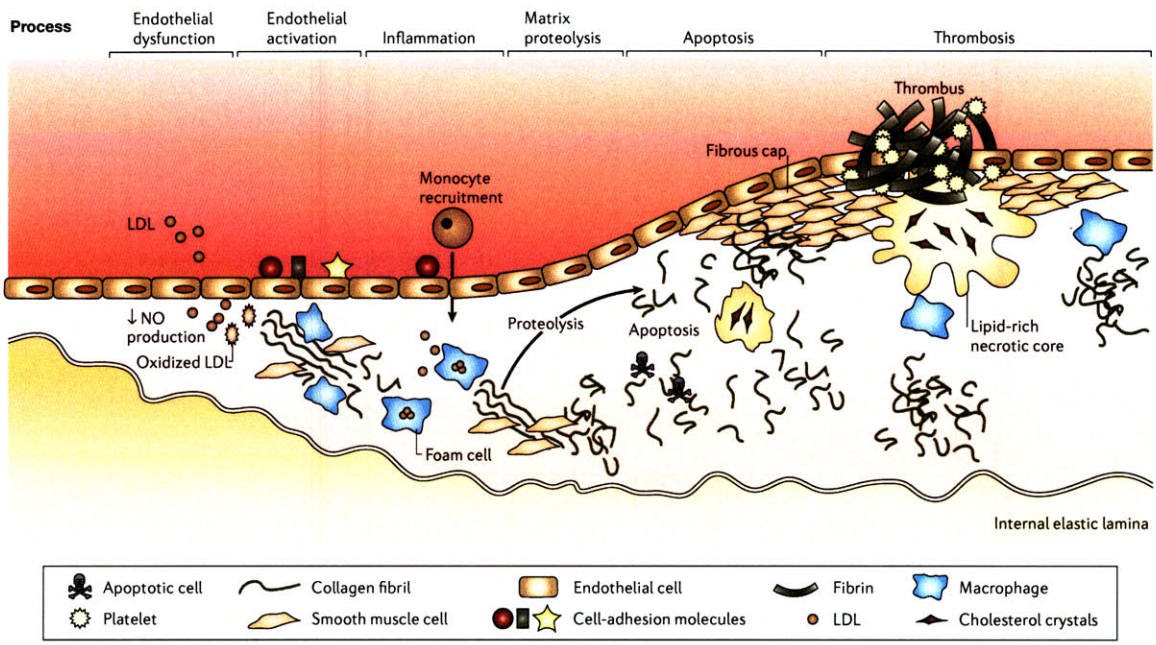


Figure 1-1: Cellular processes in the development of an atherosclerotic lesion. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics [5], © 2006.

of smooth muscle cells, embedded in a matrix of elastin and collagen. The outermost layer is the adventitia, consisting of mostly collagen [3].

1.1.1 Pathogenesis

Atherosclerosis is characterized by an accumulation of lipid and fibrous tissue in the arterial wall. The cellular processes involved in the development of an atherosclerotic lesion are illustrated in Figure 1-1. Normal endothelium is resistant to adhesion by leukocytes, but triggered by risk factors such as smoking, poor diet, hypertension, hyperglycemia, and obesity, endothelial dysfunction may occur, in which endothelial cells allow low density lipoprotein to enter into the intima. This in turn stimulates cell-adhesion molecules, which recruit monocytes to adhere to the endothelial wall and migrate into the intima [3–5].

Inside the intima, monocytes mature into macrophages and begin to engulf modified lipoproteins. Lipid droplets, primarily cholesterol esters, accumulate in the cytoplasm, and the lipid-laden macrophage is then termed a "foam cell", due to its foamy appearance when viewed under a microscope [3, 4]. This type of lesion is called a fatty streak, and is characterized by the presence of macrophage foam cells, lipid droplets in the smooth muscle cells, and T lymphocytes [4, 6].

Macrophages in the fatty streak release growth factors that stimulate migration and proliferation of smooth muscle cells into the lesion. As these smooth muscle cells accumulate in the plaque, they synthesize extracellular matrix, leading to the development of a fibrous cap [7]. At this point, lesion progression is largely regulated by inflammatory processes. Inflammatory cytokines can induce macrophages to undergo apoptosis, creating the cellular debris that forms part of the necrotic core characteristic of more advanced lesions. Mediated by the inflammation response, lipid infiltration and macrophage accumulation continue, contributing to the soft lipid in the necrotic core. These macrophages express proteolytic enzymes that break down the fibrous cap. This process can ultimately lead to the development of a lesion consisting of a lipid-rich necrotic core covered by a thin fibrous cap [2].

Alternatively, clinical observations have shown that atherosclerotic lesions develop discontinuously [8]. It is thought physical disruption of a plaque and the resulting thrombosis leads to a sudden expansion of the lesion. Three mechanisms for plaque disruption have been identified: superficial erosion of endothelial cells, microhemorrhaging within the plaque, and fracture of the plaque cap [9].

Superficial erosion is characterized by microscopic areas that lack functional endothelium. This exposes subendothelial collagen and von Willebrand factor, and ultimately leads to thrombus formation. Superficial erosion occurs commonly and is usually asymptotic [3].

Disruption of microvessels in the atheromatous lesion may also contribute to plaque progression. Lesions may develop microvascular channels as a result of neoangiogenesis. These microvessels are fragile and prone to microhemorrhage, which leads to thrombosis *in situ*. This, in turn, leads to thrombin generation, which can stimulate smooth muscle cell migration and proliferation. Additionally, activated platelets release growth factors that stimulate more collagen synthesis by smooth muscle cells, potentially leading to a growth spurt of the plaque [3].

The most common mechanism for plaque disruption is fracture of the fibrous cap. Activated macrophages within the lesion can produce proteolytic enzymes that degrade the collagen in the fibrous cap, which then weakens and becomes susceptible to rupture. Fissure of the cap allows thrombogenic materials in the lipid core to come into contact with coagulation factors in the blood, resulting in thrombus formation [4].

1.1.2 Stages of atherosclerosis

Once the initial atherosclerotic lesion has developed, it can progress discontinuously to several different stages (Figure 1-2). In an early atheroma, the intima is thickened by a lipid core covered by a fibrous cap. Often in these lesions, the vessel has expanded outward to maintain a relatively constant diameter lumen. Through a combination of the processes described in Section 1.1.1, this lesion can become a 'stabilized' plaque, with a small lipid pool and a thick fibrous cap. Alternatively, the early atheroma can progress to a 'vulnerable' plaque, characterized by a large lipid core, a thin fibrous cap, and inflammatory cells [3,5]. Rupture of the vulnerable plaque may lead to thrombosis and subsequent healing, creating a thickened fibrous intima with a narrow lumen. Rupture may alternatively lead to occlusive thrombosis, resulting in myocardial infarction.

1.1.3 Plaque stability

Our understanding of human coronary artery disease is in part derived from autopsy studies of coronary plaques from patients who have died of heart attacks [10]. These studies have reported evidence suggesting correlations between a plaque's histologic features and the presence of thrombosis. Tissue sections obtained at culprit coronary thrombi show three types of lesions: 1) the thin-capped fibroatheroma (TCFA), comprising a macrophage-rich, thin fibrous cap overlying a core of necrotic debris [11,12], 2) the endothelial erosion [10,13, 14], which may be more common in younger patients, lacks functional endothelium, and may contain higher concentrations of certain proteoglycans [15], and 3) the superficial calcific nodule [10]. The distribution of the different types of plaques implicated in myocardial infarction ranges from approximately 55-60% for TCFAs, 30-35% for erosions, and 3-7% for calcific nodules [10].

Due to its relatively high prevalence, we focus on the TCFA, the most common of these vulnerable plaques [16]. TCFAs exhibit a thin fibrous cap ($< 65 - 150$ μm) that contains smooth muscle cells in a collagenous-proteoglycan matrix, infiltrated by macrophages and lymphocytes. The fibrous cap overlies a lipid-rich core, comprised of large amounts of extracellular lipid, cholesterol crystals, and necrotic debris. It is hypothesized that thin, mechanically unstable fibrous caps may rupture when exposed to biomechanical stress, exposing the large lipid-rich core to blood. Conversely, stable plaques exhibit thicker, more

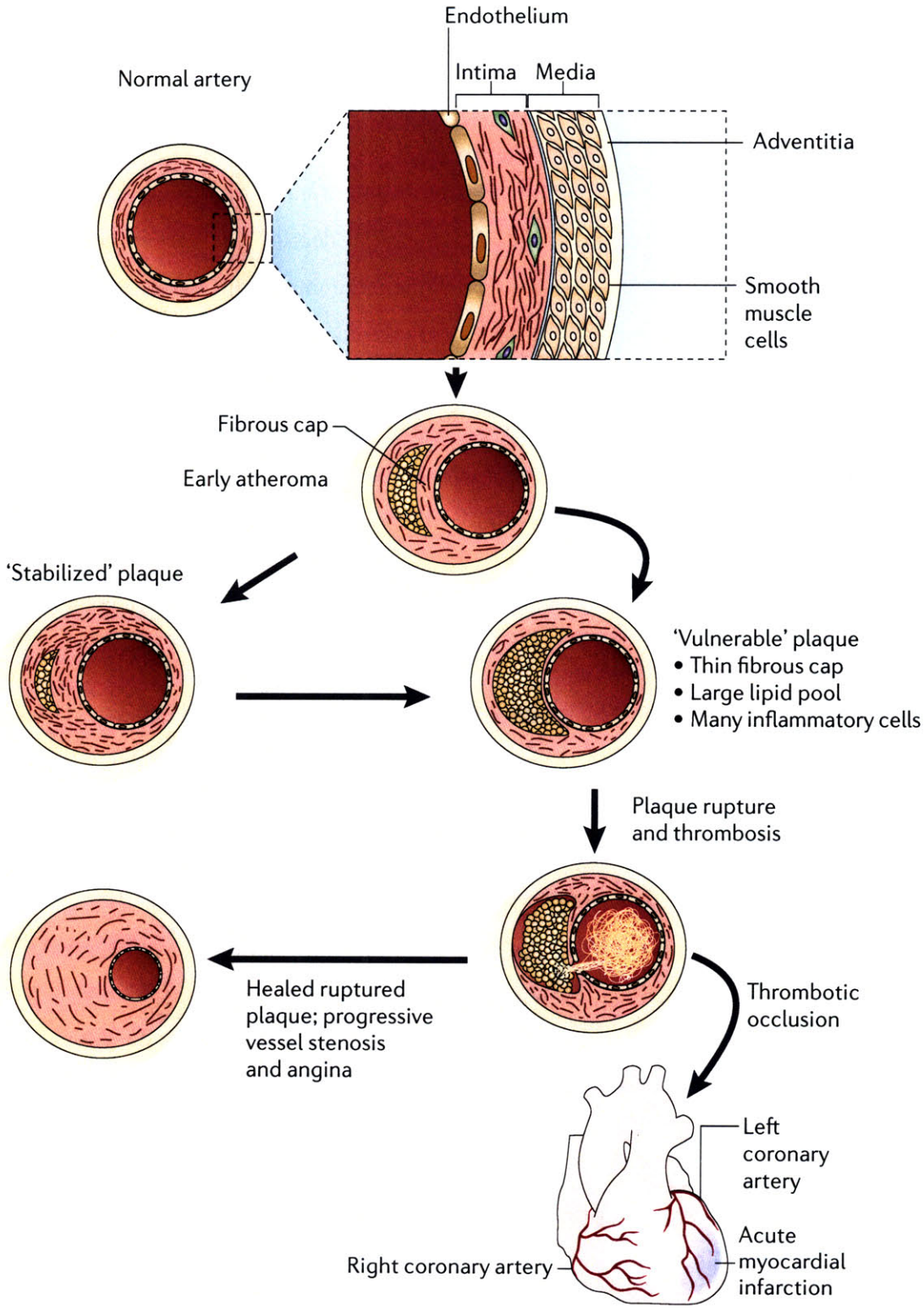


Figure 1-2: Stages of atherosclerosis. Reprinted by permission from Macmillan Publishers Ltd: Nature Reviews Genetics [5], © 2006.

robust fibrous caps and a high density of collagen and smooth muscle cells [11].

Although these autopsy studies have been valuable for determining histopathologic plaque types associated with coronary thrombosis, it is not well understood which plaques cause coronary events, why these plaques cause events, or when. For instance, TCFAs are a frequent autopsy finding in aysmptomatic or stable patients and are found with equal frequency in culprit and non-culprit arteries in acute coronary syndromes [17]. Moreover, disrupted TCFAs have been found in 10% of non-cardiac deaths [17]. Recent findings of multiple ruptured plaques [18] and increased systemic inflammation in acute patients [19] has challenged the notion of a single plaque type as the precursor for acute myocardial infarction (AMI) [20,21]. The role of erosions is also somewhat controversial, with some believing that reports on the relatively high prevalence of erosive plaques is a result of study population bias, and differing opinions on the inflammatory nature of these lesions [13, 15, 22].

While vulnerable plaque remains an important hypothesis for explaining the pathogenesis of acute coronary thrombosis, it is clear that there is much we still do not understand. For instance, in a single patient, many nonfatal plaque ruptures can occur without causing death or even symptoms [17]. Additionally, in one study of patients who experienced sudden coronary death, only 60% had lesions with plaque rupture; the remaining 40% showed superficial erosion [13]. High levels of C-reactive protein found in patients with acute coronary events suggest that systemic and multi-focal arterial inflammation may be as important as focal, vulnerable plaque features [23]. A body of evidence is growing to suggest that plaque risk is multifactorial [21]. Lesion microstructure, molecular and chemical composition, biomechanical properties, and functional state are now all considered to be important [2, 3, 7, 9, 24, 25]. Unfortunately, these factors have been difficult to prospectively study due to a lack of appropriate technologies for investigating them in patients. The development of diagnostic technologies for studying these factors, including their clinical significance and temporal evolution, is therefore merited.

1.1.4 Diagnostic modalities for atherosclerosis

Many technologies have been developed for examining atherosclerotic lesions [26,27]. While non-invasive imaging methods, such as computed tomography (CT) [28–30] and magnetic resonance imaging (MRI) [31, 32] are continually improving, their resolution and capability to determine plaque composition is limited. Angiography, the current gold standard in

clinical care, can measure lumen diameter, area, and flow resistance, but is insensitive to changes in plaque burden that are masked by arterial remodeling [26]. Angioscopy, which uses a white light endoscope to view the surface of the artery, enables detection of plaque surface features such as color, but is unable to give any information about plaque composition [33]. Cardiac magnetic resonance (CMR) non-invasively measures plaque area and volume but has limited spatial resolution [26, 27]. Intravascular ultrasound (IVUS) allows structural interrogation of the arterial wall at a resolution of about 100 μm [26, 27]. Elastography is an image processing technique based on IVUS that provides a map of the mechanical strain behavior of a lesion [34]. Electron beam computed tomography (EBCT) measures calcium content [27]. Thermography measures temperature variations at the surface of the arterial wall, under the hypothesis that heat from inflammatory cells increases temperature locally at areas of vulnerable plaques [35]. Intravascular ultrasound (IVUS) [36–39] and its more recent incarnation, Virtual Histology [40], can identify the presence of calcium. However, their capability to reliably identify lipid pools is currently under debate. Molecular contrast agents for MRI [41–44] and IVUS [45, 46] are on the horizon, but progress towards clinical studies with these techniques is encumbered by the lengthy approval process associated with new exogenous diagnostic agents.

Optical coherence tomography (OCT) is conceptually similar to IVUS, but it utilizes infrared light, which enables higher resolution images (~10 μm) of the arterial wall [47]. Extensive *ex vivo* validation studies have shown that OCT can be used to determine the presence of fibrous, lipid-rich, and calcified plaques [47–49]. OCT's strength is that it can obtain microscopic images of the arterial wall, which provides the opportunity to visualize structural features thought to be associated with high-risk plaques, such as macrophages, thin fibrous caps, and lipid pools [47, 49–52]. While OCT is excellent for characterizing plaque structure, it does not offer much direct compositional information.

The capability of investigating coronary plaques on the chemical and molecular level is needed to gain a deeper understand of CHD. Spectroscopic techniques are attractive because they have the ability to identify specific chemicals *in vivo* via optical fiber probes. Molecular imaging using fluorescent labels is a very promising technique [42, 53, 54] but faces the challenge of obtaining regulatory approval for new optical contrast agents. Autofluorescence spectroscopy and time-resolved lifetime fluorescence spectroscopy are sensitive to the presence of naturally occurring fluorophores in the artery, such as elastin, collagen, NADH, and

ceroid, but are limited by the small number of fluorophores in the artery and strong absorption in blood, which distorts the fluorescence spectrum [55–60]. Reflectance spectroscopy is sensitive to chromophores in the arterial wall, such as hemoglobin and β-carotene [61], providing the basis for "yellow plaques" seen during angioscopy [62–64]. Vibrational spectroscopy techniques provide a means of probing the detailed energy levels of molecules in the arterial wall. One such technique, near-infrared (NIR) spectroscopy analyzes absorption to determine the chemical composition of the artery. Near infrared absorption spectroscopy (NIR) has been demonstrated *in vivo*, but the broad nature of its spectroscopic features currently limits its detection to lipid-containing regions within the intima [65].

Raman spectroscopy is based on spectral analysis of inelastically scattered photons [66]. Raman spectroscopy has several advantages over NIR absorption spectroscopy. Raman spectra contain very sharp, distinct peaks compared to the broad peaks characteristic of NIR spectra. As explained above, evidence suggests that important factors for studying atherosclerosis are the presence of a fibrous cap (collagen, macrophages, lymphocytes), necrotic core (extracellular lipid, cholesterol crystals, necrotic debris), smooth muscle cells, and calcifications. It has been demonstrated that Raman spectroscopy is highly sensitive to detecting the majority of these components [67–70]. Raman spectroscopy may be an ideal tool for improving our understanding of coronary atherosclerosis, as it may provide more information about different plaque constituents than any other diagnostic modalities proposed to date.

1.2 Raman spectroscopy

When a photon interacts with a molecule it may either be absorbed, elastically scattered, or inelastically scattered. Raman scattering is inelastic; the incident photon transfers energy to the molecule, raising the molecule from its ground state to an intermediate virtual level. As the molecule relaxes to a vibrational level with higher energy than the ground state, a photon is emitted with less energy than the incident photon. This energy difference is called the Raman shift, and it is traditionally expressed in units of wavenumber (cm^{-1}). Every molecule has several unique vibrational states, resulting in scattering of photons at several different Raman shifts; thus the spectral features of the Raman scattered light can be used to identify molecules with high specificity. Furthermore, the Raman spectrum of a

heterogeneous sample composed of several chemicals is simply a linear combination of each component chemical's Raman spectrum, with relative weighting coefficients determined by concentration and Raman scattering cross section (probability for Raman scattering) [68, 71].

Spontaneous Raman scattering is very weak – for a typical sample, about 1 in 10^{10} incident photons are Raman scattered [72]. Because inelastic scattering is a low probability event, Raman photons comprise only a small portion of the light remitted from the sample, as compared to elastically scattered light and fluorescence. In human tissue, accurate chemical characterization is challenging as the inherently low Raman signal is buried in the background of these competing signals.

1.2.1 Biomedical Raman spectroscopy

Raman spectroscopy probes the vibrational modes of a sample, which depend on both the molecular composition and molecular structure of the sample. Thus, Raman spectroscopy may be sensitive to disease induced changes in biological tissue, even before such changes are histologically discernible.

As a nondestructive optical technique, Raman spectroscopy requires minimal sample preparation – the only requirement is to bring the sample close enough to the optical apparatus. The use of optical fiber probes makes it possible to conduct Raman spectroscopy on many biological tissues *in situ*, and for some tissues, non-invasively. Because of these advantages, Raman spectroscopy has been studied for the detection and diagnosis of many diseases [73], including skin cancer [74], breast cancer [75], cervical cancer [76, 77], Barrett's esophagus [78], bladder and prostate cancer [79], brain tumors [80], and atherosclerosis [69, 70, 81–85].

Instrumentation

In a typical dispersive Raman system (Figure 1-3), near-infrared laser light is focused onto the tissue. Remitted light is collected by the focusing optics and filtered to remove the elastically scattered light. The resulting autofluorescence and Raman scattered light is directed into a spectrometer, where it is dispersed onto a CCD, and the spectrum is recorded.

Because the energy shift associated with Raman scattering is independent of the wavelength of the excitation light, it is possible to choose an excitation wavelength that produces

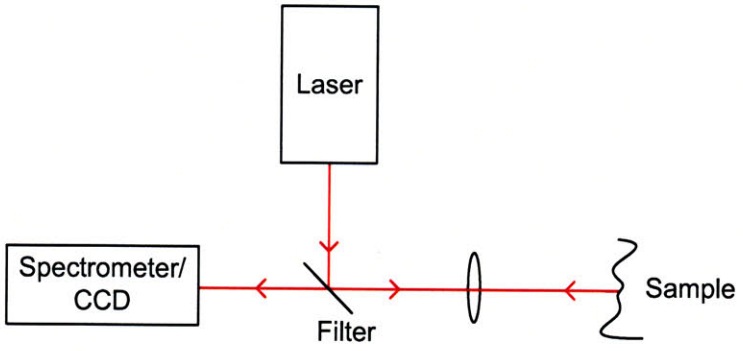


Figure 1-3: In a typical benchtop Raman system, laser illumination is reflected off a longpass optical filter and focused onto the sample. Remitted light propagates through the lens to the filter, which blocks the elastically scattered light from propagating to the spectrometer.

less fluorescence, maximizes tissue penetration, and minimizes tissue damage. Autofluorescence from the tissue can be orders of magnitude higher than the Raman signal. Near IR excitation can minimize autofluorescence, so biological Raman spectroscopy is often performed using excitation in the $700 - 900$ nm range.

Raman shifts are typically measured in the wavenumber region from 400 to 1800 cm^{-1} [68, 69, 81, 83, 86, 87], which is termed the "fingerprint" (FP) region because of its high sensitivity for biologically relevant molecules. There has been recent interest in Raman spectroscopy based on the "high wavenumber" (HW) region, from 2400 to 3800 cm^{-1}, which may contain less molecular information but offers significant technical advantages [88].

Optical fiber Raman probes

Conducting biological Raman spectroscopy *in situ* often requires a Raman probe. A Raman probe consists of one or more optical fibers, which guide excitation light from the laser to the tissue and deliver Raman scattered light from the tissue to the spectrometer. The distal tip of the optical fiber probe may contain optical elements that filter, focus, and collect the light from the tissue.

A typical FP Raman probe utilizes separate excitation and collection fibers to avoid generating a competing Raman background signal from within the fibers (Figure 1-4). Laser light propagating through a fused silica excitation fiber generates a large Raman background signal, so that both excitation light and fiber Raman are incident on the tissue. The tissue then remits the tissue Raman photons, the elastically scattered excitation photons, and the

elastically scattered fiber Raman photons. As these three signals propagate through the collection fiber, an additional fiber Raman component is generated. At the spectrometer input, the elastically scattered excitation photons can be filtered out, but the fiber and tissue Raman signals occupy the same wavelength range, and thus both contribute to the overall measured Raman signal. The fiber Raman signal often has spectral features that overlap with the tissue Raman signal, and in most cases the fiber Raman signal is much higher in magnitude than the tissue Raman signal.

This background can be avoided by placing optical filters at the distal end of the fibers [86,89,90]. On the excitation fiber(s), a narrow bandpass or shortpass filter is used to block the fiber Raman signal from entering the tissue, so that only excitation light impinges on the tissue sample. A longpass filter is placed on the collection fiber(s) to block elastically scattered light remitted from the sample, allowing only Raman scattered light from the tissue to propagate through the fiber(s) to the spectrometer.

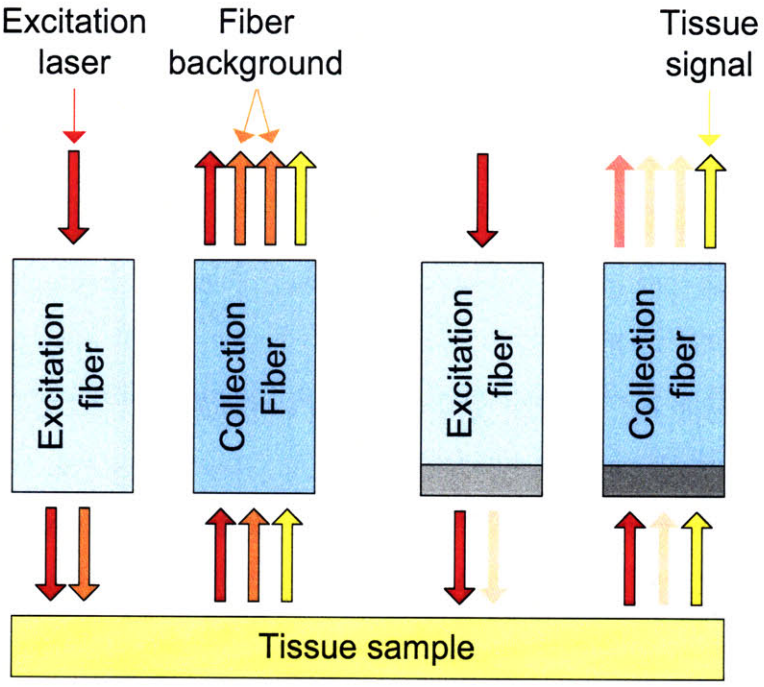


Figure 1-4: Raman probes for the fingerprint region require optical filtering to avoid confounding the biological Raman signal with the Raman signal generated within the fiber.

Data analysis and classification

Biomedical Raman spectroscopy applications have primarily been focused on disease diagnosis, which requires the development of a diagnostic algorithm. The algorithm provides a mapping from the multi-dimensional Raman spectrum to one of a small number (< 10) of distinct classes, or diagnoses. Raman spectroscopy diagnostic algorithms are often performed in two steps: dimension reduction and classification.

The measured Raman spectrum may consist of ~1000 data points (corresponding to the number of CCD pixels), so for further analysis it is convenient to reduce the dimension of the data to a tractable number of parameters. Simple examples of dimension reduction techniques include extracting the intensities of a carefully chosen subset of Raman bands, or calculating ratios of specific Raman bands [77]. Another method of dimension reduction is to model the Raman spectrum as a linear combination of a limited number of basis spectra. The set of bases used in the model can be physically motivated, such that they correspond to spectra of specific chemical [68, 70] or morphological constituents [83, 91], or they may be a numerically derived set of orthogonal bases, as in principal components analysis [81, 92, 93]. When using physical bases, the weighting coefficients are found using a fitting algorithm (typically a least squares algorithm), but for orthogonal bases, weighting coefficients are found by simply computing the dot product of each basis with the spectrum.

Next, the output parameters from the dimension reduction step are used as the input to a classification algorithm. Supervised classification is often used, requiring a training set consisting of Raman spectra and corresponding disease diagnoses. Examples of classification algorithms that have been previously demonstrated with Raman spectroscopy are linear discriminant analysis [93], logistic regression [69, 70, 81], and artificial neural nets [92].

1.2.2 Raman for atherosclerosis

Several researchers have investigated the use of Raman spectroscopy for diagnosing atherosclerosis [68–70, 81–87, 91, 92]. Two physically motivated spectral models have been developed. The first model utilized basis spectra measured from purified chemicals found in the artery, including delipidized artery (primarily collagen and elastin), free cholesterol, cholesterol ester, triglycerides, calcification, and β-carotene [68]. For the second spectral model, a confocal Raman setup was used to acquire basis spectra of morphological structures found

Table 1.1: Several studies utilizing Raman spectroscopy for diagnosis of atherosclerotic plaque have been reported. Raman spectra are modeled as a linear combination of chemical bases (chem.) [68], morphological bases (morph.) [91], or principal components (PCA), or reduced using the discrete wavelet transform (DWT). Where available, the number of specimens and classification accuracy for the training and prospective data sets is given.

			Training set		Prospective set	
Author	Artery	Model	N	Accuracy	N	Accuracy
Römer [70]	coronary	chem.	97	98%	68	94%
Deinum [81]	coronary	PCA	97	94%	68	90%
Buschman [69]	coronary	morph.	97	98%	68	94%
Salenius [82]	carotid/femoral	chem.	167	—[a]	0	—
Motz [83]	carotid/femoral	morph.	34	82%	0	—
Silveira [84]	coronary	PCA	60	88%	51	82%
Nogueira [85]	carotid	PCA	75	94%	0	—
de Paula [92]	coronary	PCA DWT	40	> 93% > 91%	20	> 92% > 89%

[a]The authors do not report overall accuracy. The sensitivity and specificity for detecting lipid was 92% and 53%, and for detecting calcium salts, 100% and 88%.

in the artery, including internal elastic lamina, collagen fibers, smooth muscle cells, adventitial fat, foam cells, cholesterol crystals, carotene containing crystals, and calcifications [91]. These models, in addition to models numerically derived using principal component analysis (PCA), have been used to diagnose atherosclerotic lesions via Raman spectroscopy (Table 1.1).

In three related studies, a single dataset was analyzed using all three spectral models. Fingerprint Raman spectra were acquired from a single site on 165 *ex vivo* coronary lesions using a benchtop configuration. Lesions were irradiated with 350 mW of 830 nm laser light, and Raman spectra were collected for between 10 to 100 seconds. The spectral acquisition site was marked with ink and submitted for histological processing. Each specimen was histologically classified into one of three general categories: nonatherosclerotic plaque (73 specimens in the training set and 26 in the prospective set), noncalcified atherosclerosis (9 in training set and 16 in prospective set), and calcified atherosclerotic plaque (15 in training set and 26 in prospective set).

Based on this dataset, Römer *et al.* fit the spectra to the purified chemical based spectral model, and using the resulting fit coefficients, diagnosed the lesions using logistic

regression. They report 98% accuracy in the training dataset and 94% accuracy in the prospective dataset [70]. Deinum *et al.* used the same data set to derive basis spectra using PCA. Fit coefficients for two principal components served as the input to logistic regression, which yielded accuracy of 94% and 90% for the training and prospective datasets, respectively [81]. Buschman *et al.* fit the dataset to the morphological spectral model and used the fit coefficients with logistic regression to acheive accuracy of 98% and 94% for the training and prospective datasets.

In a separate study, Salenius *et al.* acquired spectra of 167 carotid and femoral artery lesions. After fitting the spectra to the chemical spectral model, the presence of diseased levels of cholesterol and calcium salts were detecting using threshold tests on the fit coefficients. They reported 92% sensitivity and 53% specificity for detecting lipid, and 100% sensitivity and 88% specificity for detecting calcium salts [82].

Motz *et al.* performed Raman spectroscopy *in vivo*, acquiring 34 spectra during carotid endarterectomy and femoral bypass surgeries. The spectra were fit to the morphological model, and logistic regression was used to classify plaques as vulnerable or stable. Histologically, vulnerability was assessed using a scheme proposed by the authors, in which a vulnerability score was assigned based on histologic features such as the presence of thrombosis, ulceration, and hemorrhage, the thickness of the fibrous cap, the diameter of the necrotic core, and the presence and depth of foam cells. Using leave-one-out cross-validation, the authors report sensitivity and specificity of 79% and 85% for detecting vulnerable plaque, with an overall accuracy of 82% [83].

Silveira *et al.* acquired Raman spectra at single sites of 111 coronary artery specimens and histologically classified the specimens into three categories: nonatherosclerotic, noncalcified, and calcified. They derived a spectral model using PCA, and using discriminant analysis based on the Mahalanobis distance, two principal components were used to classify the specimens, with an overall accuracy of 88% in the training set and 82% in the prospective set. They report sensitivity and specificity of 90% and 83% for prospectively detecting lipid rich, noncalcified plaques ($n = 21$) [84].

Nogueira *et al.* performed a similar study, acquiring spectra at single sites of 75 carotid artery specimens and histologicaly classifying the specimens into the same three categories. Again, two principal components were used to classify plaques using discriminant analysis based on the Mahalanobis distance. They report overall accuracy of 94%, with sensitivity

and specificity of 89% and 96% for detecting lipid rich, noncalcified plaques ($n = 19$) [85].

de Paula Jr. and Sokki acquired Raman spectra at single sites of 60 coronary artery specimens, which were histologically diagnosed into three categories: atheroma, non-pathologic tissue, and calcified tissue. The authors compared two dimension reduction techniques, the discrete wavelet transform (DWT) and PCA. Between 10 and 19 coefficients were used to train and validate artificial neural networks, with overall accuracy of greater than 90% for both DWT and PCA [92].

1.2.3 Catheter-based Raman spectroscopy for atherosclerosis

Conducting intracoronary Raman spectroscopy *in vivo* requires a catheter, consisting of an optical probe placed inside a sheath that can be safely maneuvered inside the coronary artery tree Figure 1-5). The optical probe itself is made up of one or more bundled optical fibers that guide light to and from the distal optical elements that filter, focus, and collect the light from the tissue. Towards the proximal end of the catheter, the fiber bundle is split into excitation and collection fibers. The excitation fiber is coupled to the laser and the collection fibers are coupled into the input of the spectrometer.

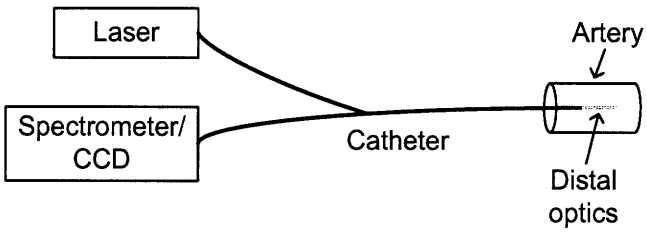


Figure 1-5: A Raman catheter consists of one or more optical fibers to: 1) deliver laser light to the artery, 2) collect emitted Raman light, potentially using distal optics, and 3) deliver the collected Raman light to the spectrometer.

Several FP Raman probes have been developed for Raman spectroscopy of arterial tissue (Table 1.2) [86, 94, 95]. All of these probes consist of a central excitation fiber surrounded by a ring of collection fibers, with separate filters used for excitation and collection. The fibers used are generally multimode fused silica fibers.

Shim *et al.* demonstrated a 1.5 mm diameter probe consisting of seven collection fibers (300 μm core diameter) arranged around a single excitation fiber (400 μm core). Each fiber was internally filtered with an individual filter placed 2.5 cm from the distal tip of the probe. The collection fibers were beveled and coated with a reflective layer to increase

the overlap between the excitation and collection light cones. This forward-viewing probe was demonstrated on rabbit aorta *ex vivo* [94]. A side-viewing version of this probe was made, without beveled fibers, by incorporating a gold plated mirror oriented at 45° from the optical axis. Raman spectra were acquired with the side-viewing probe *in vivo* in sheep aorta [87].

Motz *et al.* fabricated a 2 mm diameter Raman probe consisting of 15 collection fibers arranged around a single excitation fiber. All fibers had 200 μm core diameter. Two separate optical filters were epoxied to the distal tip of the fibers: a central cylindrical shortpass filter for the excitation fiber and an annular longpass filter for the collection fiber. A 2 mm diameter ball lens was incorporated at the distal tip to maximize overlap of the excitation and collection light cones [86]. This forward-looking probe was used in *in vivo* during carotid endarterectomy and femoral bypass surgeries [83].

Komachi *et al.* fabricated a 0.6 mm diameter Raman probe consisting of 9 fibers with 114 μm core diameter, with 8 collection fibers arranged around a single excitation fiber. Similar to the probe demonstrated by Motz *et al.*, one cylindrical and one annular filter were epoxied to the optical fibers. No modifications were made at the distal end of the probe to modify the excitation or collection light cones. The forward looking probe was demonstrated on rabbit aorta *ex vivo* [95]. Using a 45° mirror, the Raman probe was incorporated into a 2 mm diameter side-viewing catheter, which was used to measure Raman spectra of a lesion-mimicking phantom [96]. They have recently reported a modified probe, incorporating a lens with a hole drilled through the center, providing for lensing over the the collection fibers only [97]. To date, this 0.7 mm diameter probe has not been demonstrated on arterial tissue.

1.3 Toward an intracoronary Raman system

A successful *in vivo* Raman spectroscopy system must achieve sufficient signal-to-noise ratio (SNR) for accurate data analysis and classification while utilizing a small (~1 mm) diameter catheter. It must also have the ability to screen a large area of the coronary at high speeds (< 1 s per point), be portable, and meet patient safety standards. The SNR is defined as the ratio of the mean of a particular Raman band to its standard deviation, and is both sample and system dependent. SNR is affected by 1) the magnitude of the Raman

Table 1.2: Several optical fiber Raman probes for arterial tissue have been reported.

	Shim [94]	Motz [86]	Komachi [95]
Fibers	7 around 1 (400 μm core ex. 300 μm core coll.)	15 around 1 (200 μm core)	8 around 1 (114 μm core)
Filter	1 filter per fiber	1 central filter 1 annular filter	1 central filter 1 annular filter
Lensing	beveled collection fiber tips	ball lens	none
Outer diameter	1.5 mm	2 mm	0.6 mm
Applications	*In vivo* sheep aorta [87] 830 nm, 100 mW 30s	*In vivo* human femoral and carotid (surgical field) [83] 830 nm, 100 mW, 1s	*Ex vivo* rabbit aorta [95] 720 nm, 7 mW, ~20 s

signal, 2) the collection efficiency of the optical fiber probe, 3) the magnitude of the Raman background generated in the probe, and 4) the presence of shot noise due to the relatively large fluorescence signal generated in the sample. Increasing the laser power or exposure time will increase the total laser energy incident on the sample, with a corresponding increase in the Raman signal. However there is an upper limit to the amount of energy than can be safely delivered to the patient.

High wavenumber Raman

Several of these challenges may be addressed by implementing a significant change in the basic Raman system. In this thesis, we investigate Raman spectroscopy in the high wavenumber (HW) region, which provides three important advantages. First, Raman signals in the HW region are roughly three times greater than those in the FP region [88]. Second, very little fluorescence is generated in this region, reducing the overall shot noise. Finally, fused silica does not have Raman bands in the HW region. As a result, fiber background is minimal and filters are not needed, simplifying catheter fabrication [88, 98].

These factors indicate that HW Raman may result in increased SNR, allowing for faster data acquisition. However, HW Raman spectroscopy detects different molecular vibrational states than are seen in the FP region. An unpublished proof-of-principle study of HW Ra-

man in five coronary artery samples found that while calcification cannot be seen in the HW region (2700 to 3100 cm^{-1}), there is high sensitivity to triglycerides, cholesterol, cholesterol esters, and proteins [88]. Due to the significant technical advantages of HW Raman spectroscopy, a comprehensive investigation of its capabilities for chemical characterization is merited.

One of the goals of this thesis is to develop a database of FP and HW Raman spectra with correlated histology. In Chapter 2 we describe the methods used to create this database, and in Chapter 3 we compare the diagnostic capability of Raman spectroscopy in the fingerprint and high wavenumber regions.

Raman line imaging

Previously published work on Raman spectroscopy for atherosclerosis has concentrated on assigning a diagnosis based on the Raman spectrum taken from a single site within a lesion. However, this approach is limited because atherosclerotic lesions are complex and heterogeneous, and a single site spectrum may not capture enough chemical information to fully characterize the lesion. For this reason, we investigate a Raman line imaging procedure, in which multiple Raman spectra are acquired, at fixed intervals, across the full cross-section of the lesion. This line imaging paradigm is analogous to the way we expect data will be taken with the intracoronary cathether.

By acquiring localized data across a larger area of the plaque than is traditionally used, the data from the line imaging paradigm may enhance plaque classification algorithms. In particular, it may become possible to extend the traditional 3-category classification scheme to a more sophisticated scheme, based on the chemical content and spatial extent of the entire lesion. For example, the ability to subdivide the "non-calcified plaque" category into "necrotic core" and "non-necrotic core" plaques would be clinically relevant.

We have used the Raman line imaging paradigm on all of the specimens in the combined FP/HW *ex vivo* Raman database. This paradigm has added new challenges in managing the data. We describe the line scanning methods in Chapter 2.

Optical probe design

A key component of the intracoronary Raman system is the intracoronary catheter. In this thesis we explore optimization of the optical probe design to increase collection efficiency.

Building on prior work in the field, we fabricate and characterize a prototype Raman probe in Chapter 4. In Chapter 5, we convert the forward-viewing Raman probe to a side-viewing catheter and demonstrate it in an *in vivo* xenograft model, in which human coronary arteries extracted from a cadaver are grafted onto the heart of a living swine. In Chapter 6, we describe an optical modeling technique for simulating probe sampling volume and collection efficiency, which can be used for optimizing Raman catheter design.

Finally, we have concluding remarks in Chapter 7.

Chapter 2

Ex Vivo Raman Database

In this chapter we describe the *ex vivo* Raman database, a database of Raman spectra and coresponding histology for over sixty atherosclerotic plaque specimens obtained from human aorta. Generating the database is a multiple step process that includes acquisition and processing of Raman spectra and histological images, verification and validation of the data, and pathological diagnosis (Figure 2-1). In this chapter, we describe the methods used to build and visualize the database. Data analysis will be discussed in Chapter 3.

2.1 Data acquisition

2.1.1 Experimental setup

We have developed a benchtop Raman system that implements both high wavenumber Raman spectroscopy and line imaging. Using a single fixed-grating spectrometer and CCD, the system is capable of acquiring Raman spectra in both the fingerprint (FP) ($600-1800$ cm^{-1}) and high wavenumber (HW) ($2600-3100$ cm^{-1}) regions by sequentially illuminating the sample with two separate laser sources. The system also incorporates line scanning, physiological temperature control and automated depth focusing.

The Raman line imaging system is shown in Figure 2-2. The Raman spectra are detected by a fixed-grating spectrometer (Holospec f/1.8, Kaiser Optical Systems; Ann Arbor, MI) and a deep-depleted, back illuminated CCD (Pixis 400:BR, Princeton Instruments; Trenton, NJ); this combination acquires spectra in the wavelength range of 830 to 1000 nm. When illuminating the samples with an 830 nm diode laser (Process Insruments Inc.; Salt Lake

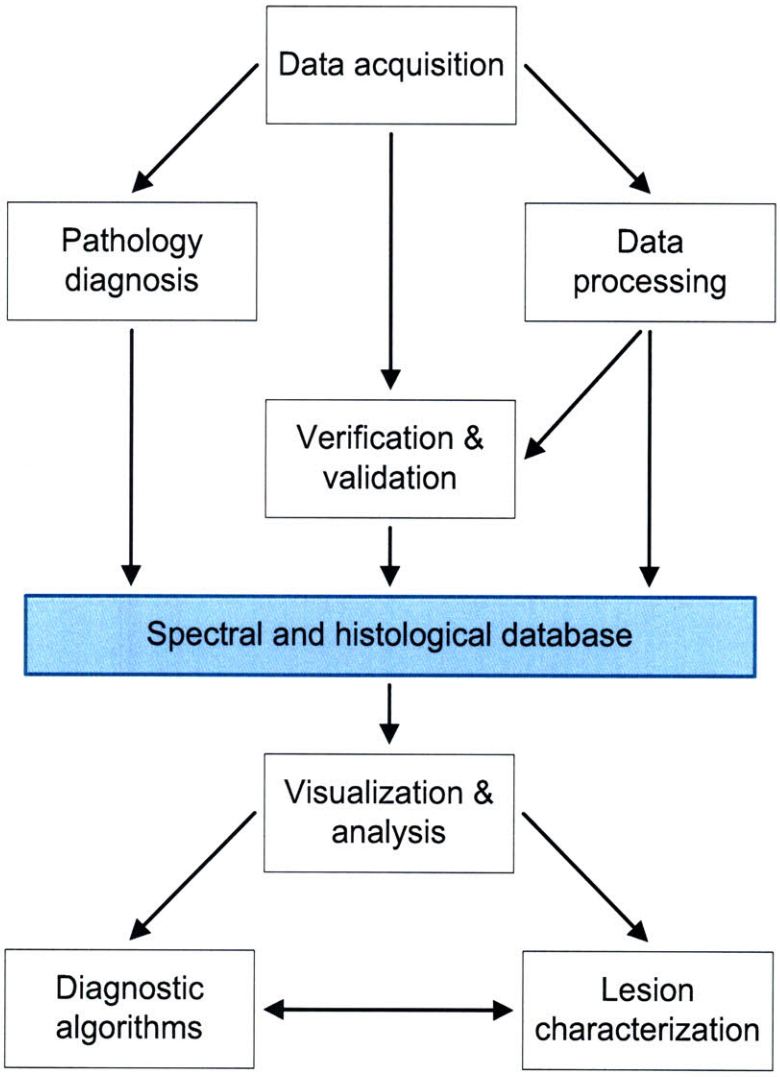


Figure 2-1: Creating the *ex vivo* Raman database.

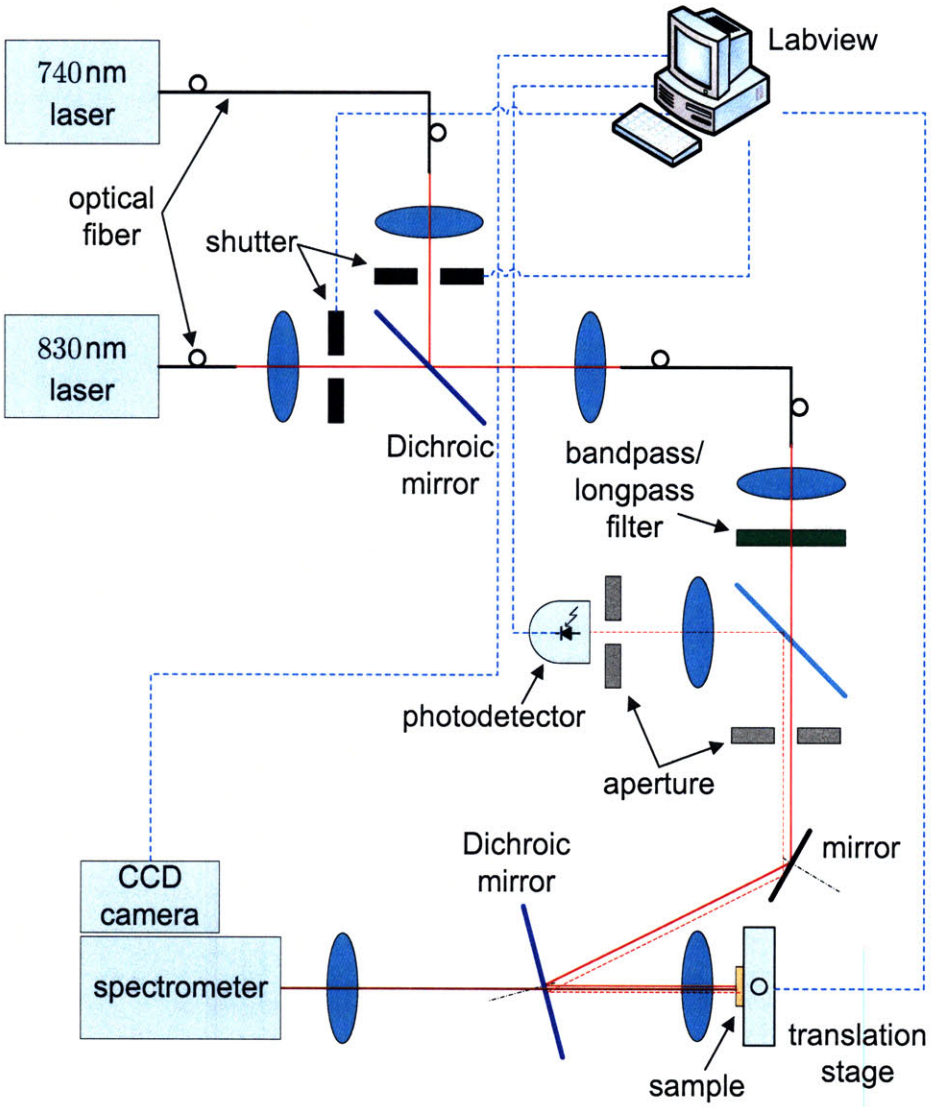


Figure 2-2: Benchtop Raman line imaging system.

City, UT), the system measures Raman spectra in the FP region (400 – 1800 cm^{-1}), and when illuminating with a 740 nm diode laser (Innovative Photonics Solutions; Monmouth Junction, NJ), HW spectra are measured (1600 – 3300 cm^{-1}).

Light from the two fiber-coupled diode lasers was collimated, combined into a single beam path using a dichroic mirror, and focused into a 100 μm core, multimode fiber with 0.26 NA (Polymicro Technologies; Phoenix, AZ). Two computer-controlled high speed mechanical shutters (Uniblitz, Vincent Associates; Rochester, NY) were placed before the dichroic mirror to rapidly switch between the two excitation wavelengths. Light exiting the common fiber was collimated and any Raman or fluorescence generated within the fiber was removed using a laser line bandpass filter (Semrock LL01-830; Rochester, NY) at 830 nm, or a longpass filter (CVI) for the 740 excitation beam. The laser light was then reflected off a longpass filter (Semrock LP02-830RU) at a ~10° incidence angle and focused on the sample using a f/2 calcium fluoride lens. Light returning from the sample was directed toward the longpass filter by the same f/2 lens. The longer wavelength Raman signal and tissue fluorescence passed through the longpass filter and were focused into the spectrometer. The Holospec spectrometer had a 100 μm internal slit resulting in an average spectral resolution of 8 cm^{-1}.

Temperature control

The aortic specimen was mounted on a custom-made specimen mount, designed to maintain the specimen at a physiological temperature of 37° C (Figure2-3). Temperature was maintained by flowing water through the mount from a thermostated water bath (NesLab, Thermo-Fisher Scientific; Waltham, MA). Prior to line imaging, each specimen was pinned to cork inserts to prevent movement during the scan, and the specimen was immersed in a phosphate-buffered saline solution to prevent dehydration.

Raman line imaging

Line imaging of each aortic lesion was implemented by securing the specimen mount to a computer controlled 3-axis translation stage. Each translation stage axis was equipped with a linear actuator (LTA-HL, Newport; Irvine, CA) driven by a computer interfaced controller (Newport ESP3000). Spectral line images were acquired in one direction (Figure 2-3). Raman spectra were acquired at discrete locations on the specimen, with a fixed

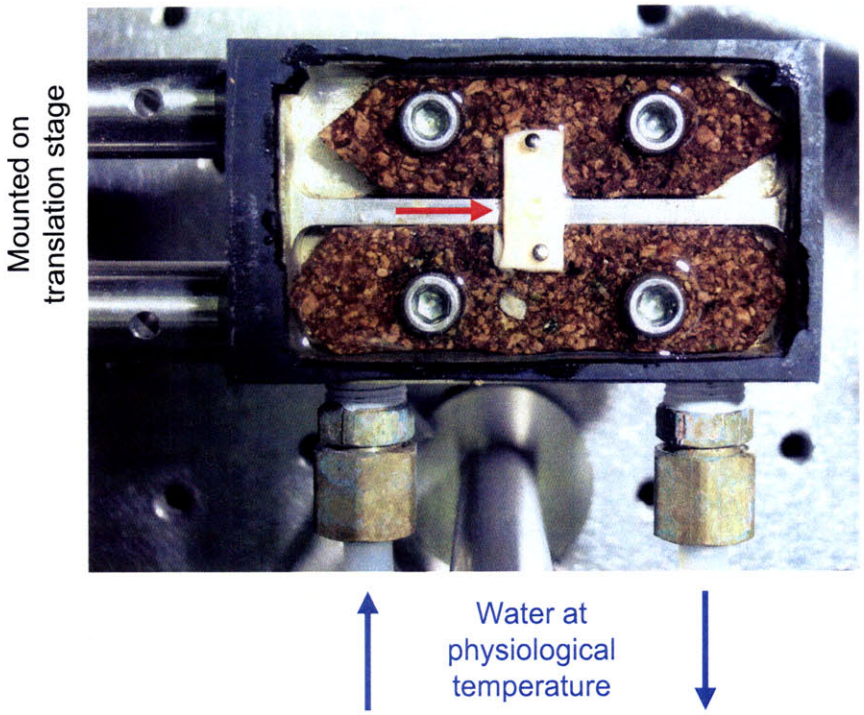


Figure 2-3: Temperature-controlled specimen mount.

interval of $200 - 250$ μm.

Automated depth focusing

The tissue surface height of an aortic lesion is typically nonuniform, so we have implemented automatic z-focusing to ensure that all spectra are taken with the same distance between the tissue surface and focusing lens, thus maintaining a constant beam diameter at the surface. Depth focusing was implemented by maximizing the amount of specularly reflected excitation light collected from the tissue surface (Figure 2-2).

The excitation light returning from the tissue was collected by the f/2 lens, collimated, and reflected off the longpass filter, back into the incoming beam path (dashed line in Figure 2-2). Since the returning excitation light is made up of diffusely and specularly reflected components, an aperture was placed in the beam path to remove some of the diffuse reflectance. The remaining light was directed toward a photodetector and focused into a 100 μm pinhole via an achromatic f/2 lens. A LabView routine adjusted the focus (z-axis) automatically to maximize the specular intensity recorded by the photodetector.

2.1.2 Spectral and histological data acquisition

The process of spectral and histological data acquisition includes tissue acquisition, plaque selection, Raman line imaging and histological processing. Detailed notes were taken through the acquisition procedures; these notes were later used for validation.

Tissue acquisition

Human aortic specimens were obtained from the Massachusetts General Hospital (MGH) Autopsy Service and the National Disease Research Interchange.[1] Arterial tissue was harvested within 48 hours after death, and aortic segments were cleaned of excess connective tissue and stored at $-80°$ C until imaging.

Lesion selection

Prior to imaging, whole aortic segments were passively thawed, rinsed in PBS and stored on wet gauze at $4°$ C until needed. Based on visual inspection, slightly raised lesions with a yellow tint were selected for inclusion in the database.

A laser speckle imaging instrument was used to help identify necrotic core lesions. Laser speckle imaging has been successfully demonstrated for identifying necrotic core plaques and can be employed as a rapid qualitative method for screening lesions believed to have necrotic cores or large lipid pools [99–101]. Although quantitative analysis of the speckle pattern was not performed, the dynamic speckle pattern observed for necrotic cores and in particular thin capped cores demodulates rapidly, in contrast to fibrous and calcified plaques, which have more static speckle patterns.

Once a lesion was identified, it was prosected from the aorta, taking note of the direction of blood flow and other landmarks, such as side branches. The specimen was then placed in the sample mount, with the region to be line imaged aligned directly over the aluminum bar (Figure 2-3). The tissue specimen was pinned to the cork inserts to prevent movement and immersed in PBS to prevent desiccation during the scan. Each specimen was assigned an identification number: UMD001, UMD002, etc.

[1] Aortic specimens were obtained with approval from the Massachusetts General Hospital Institutional Review Board (protocol #2006-P-002040/1).

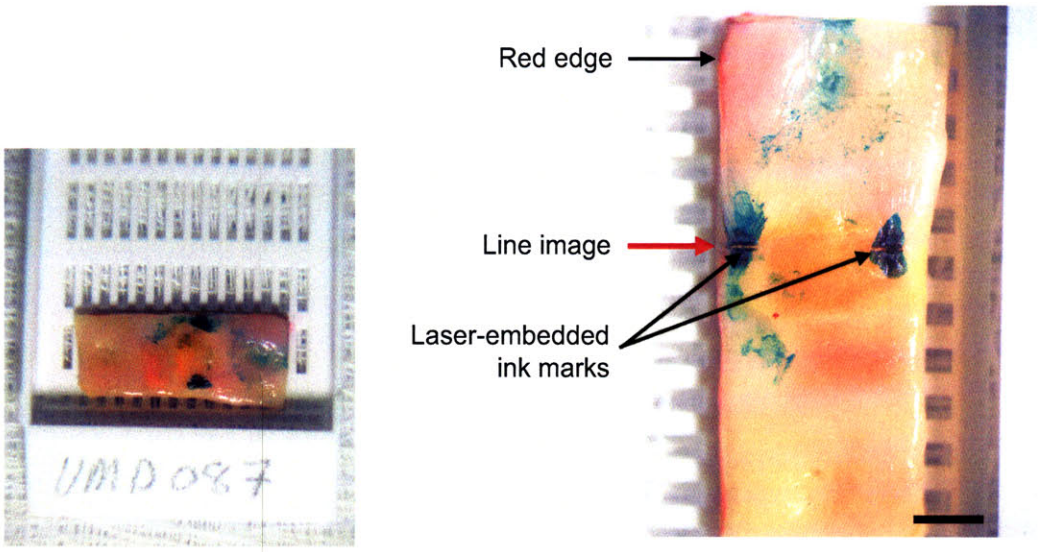
Calibration spectra

Prior to specimen imaging, instrument calibration data was acquired. For wavenumber calibration, the Raman standard acetaminophen was used, which has numerous well-defined peaks. Calibration data for spectral response correction was obtained by diffusely reflecting white light from a NIST-traceable, fiber-coupled tungsten source (Ocean Optics; Dunedin, FL) off a barium sulfate sample placed at the focus of the excitation beam. Finally, the power delivered by each laser source was verified by monitoring the signal strength of a peak in the cyclohexane Raman spectrum.

Raman line scanning

For each specimen, FP spectra were acquired with excitation power of 100 mW (830 nm), and HW spectra were acquired with excitation power of 70 mW (740 nm). Raman spectra were acquired at each site for typically 110 frames with an integration time of 0.25 seconds. Thus, each site was interrogated for 27.5 seconds before moving to the next position, 200 or 250 μm away. The start and stop positions of the scan were chosen such that the excitation beam completely scanned through the whole specimen, beginning and ending on the aluminum bar of the specimen mount (Figure 2-3). Because spectra acquired from the tissue have different bands than spectra acquired with the excitation beam off the tissue, it is easy to identify which spectra correspond to the beginning and end of the tissue, allowing accurate registration of the spectral data to the histologocial sections.

Once both HW and FP Raman scans were complete, the PBS solution was removed from the specimen holder and registration marks were burned into the lumen (Figure 2-4). For this process, green ink was placed on the tissue surface directly under the excitation beam and exposed for approximately 20 seconds, producing a small ink burn on the tissue surface. Since the laser wavelength was strongly absorbed by the ink, the registration process caused minor thermal damage to the tissue, resulting in tissue shrinkage on the edges. This can be seen grossly as a concave feature at the edges of the tissue adjacent to the registration marks (Figure 2-4). In histology sections, the tissue shrinkage appears as contraction of the lumenal surface (Figure 2-5).



(a) (b)

Figure 2-4: Registration marks in gross pathology. Scale bar, 2.5 mm. (UMD087)

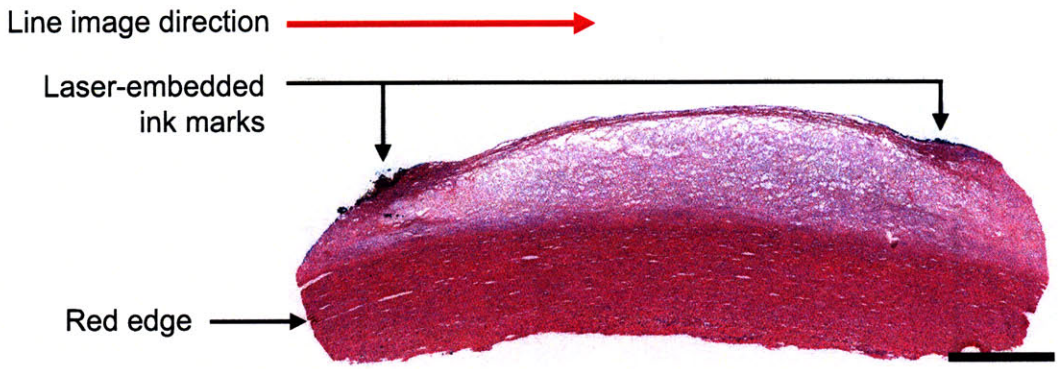
2.1.3 Histological processing

After the specimen was removed from the specimen mount, red ink was applied to the starting edge of the image scan (Figure 2-4). The specimen was then placed in a labeled tissue cassette and photographed with the specimen identification number visible (Figure 2-4(a)). A second, higher magnification photo was then taken (Figure 2-4(b)).

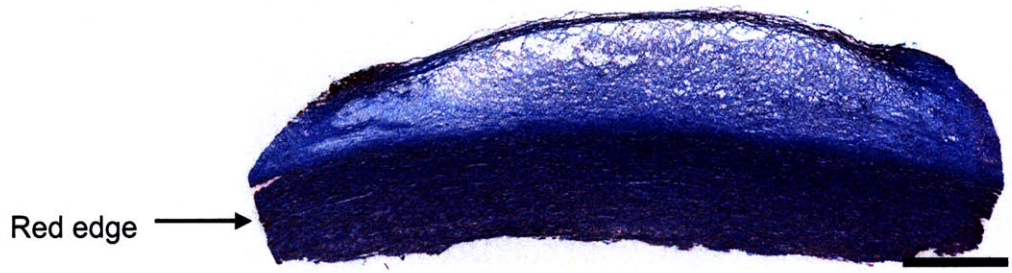
Once the gross pathology was photographically documented, the specimen was bisected across the green registration marks. Two parallel razor blades were used to cut the back tissue edge so that it was parallel with the front face. With the red edge oriented on the left, the lower bisected half was fixed in formalin, while the upper half was frozen and reserved for stains requiring fresh tissue.

All formalin-fixed specimens were placed in decalcifying solution for $2-4$ hours, after which they were rinsed in water and stored in formalin until being sent to the MGH Pathology Service for routine histological processing. The paraffin embedded specimens were cut into 5 μm sections and stained with trichrome and hematoxylin and eosin (H&E). Three additional unstained slides were prepared for future staining, as necessary.

All histology slides were digitized at a native resolution of 20X, using the Aperio slide scanning system (Aperio Technologies; Vista, CA). Digitizing the histology slides enables correlation of the Raman line image data with lesions morphology as a function of position.



(a) H&E



(b) Trichrome

Figure 2-5: Registration marks in the histological sections. Scale bar, 1 mm. (UMD087)

2.2 Data processing

Once the spectral and histological data for each specimen was acquired, several processing steps were necessary to extract the Raman spectra from the acquired tissue spectra and to register the spectra to sites on the histological image. All data processing routines were implemented in MATLAB version 7.5 (The MathWorks; Natick, MA).

2.2.1 Spectral processing

Preprocessing

Preprocessing of the acquired raw data was necessary to extract the Raman spectrum from confounding signals arising from the environment, the instrument, and fundamental noise sources. Extraneous signals that were removed or minimized through processing include cosmic rays (stray atmospheric radiation) impinging on the CCD, CCD pixel response non-uniformity, and shot noise.

Preprocessing steps are illustrated for a representative specimen (UMD087) in Figure 2-6. The topmost black spectra show the raw data, consisting of 100 frames of data taken with

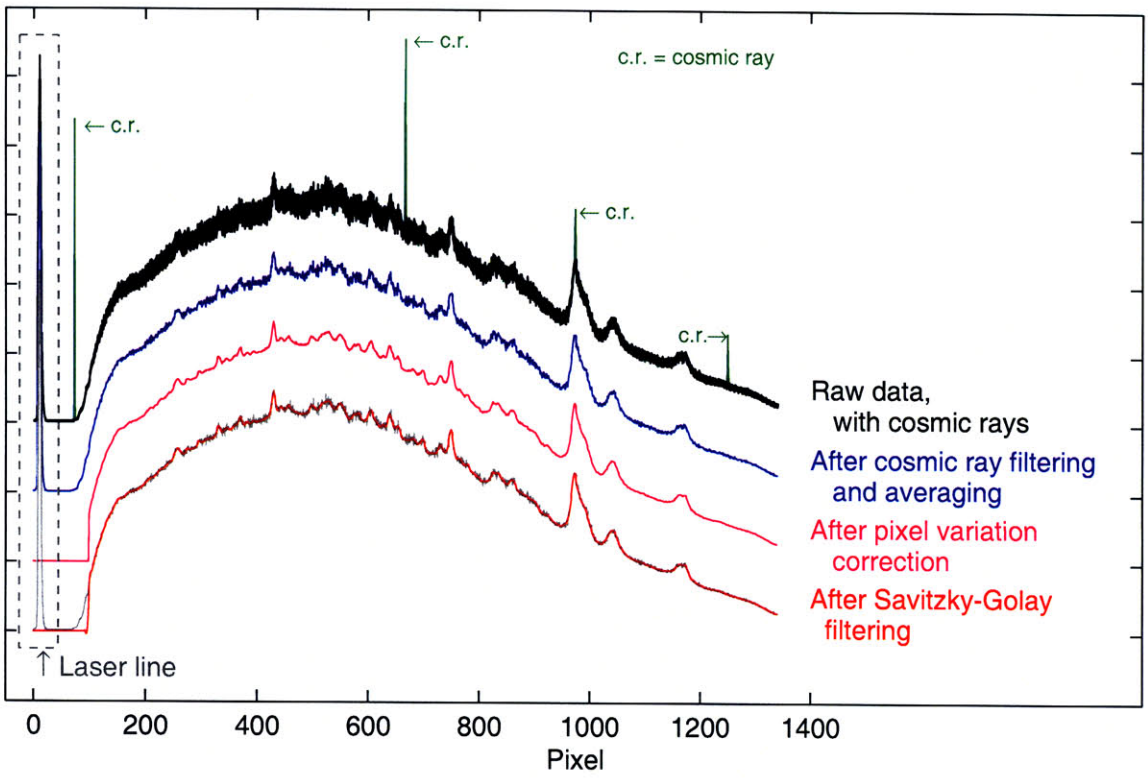


Figure 2-6: Spectral preprocessing. Of the 100 frames of raw spectral data (blue), 3 frames have spikes due to cosmic rays (green). The cosmic ray-filtered, frame-averaged data (blue) exhibits very high frequency noise due to pixel-to-pixel variations in the CCD, which was removed using a calibration factor (magenta). The resulting spectrum was filtered twice with a Savitzky-Golay filter (red; for comparison, the non-pixel corrected spectrum is underlaid in gray). The PRNU calibration factor is shown in Appendix A (Figure A-3).

an exposure time of 0.25 seconds each. In this example, the FP spectra were contaminated by four cosmic rays, shown in green. Cosmic rays were removed using a threshold-triggered temporal smoothing filter, and the resulting spectra were then averaged together to increase signal-to-noise ratio, as shown in blue. Noise due to CCD pixel response non-uniformity was removed by dividing each pixel by a calibration factor proportional to its quantum efficiency, which was estimated *a priori* for the CCD and the illumination conditions. This estimation procedure is described in detail in Appendix A. The corrected spectrum is shown in magenta. For display, the spectrum was smoothed by applying a Savitzky-Golay filter twice [102], as shown in red.

One preprocessing step that we did not perform is correction for the wavelength dependent system response. The optical components in the Raman system, such as the lenses and filters, as well as the CCD, have a wavelength-dependent response, which can result in

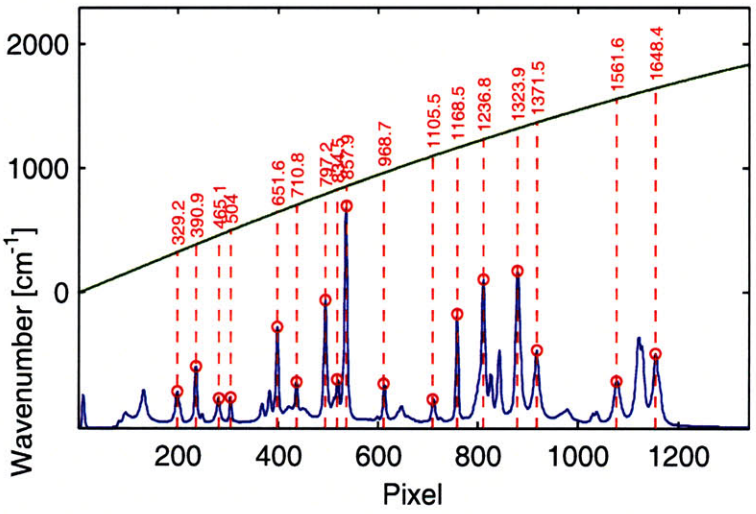
the measured relative intensity of two Raman bands being different from the actual relative intensity. One method of correcting for this wavelength-dependent system response is to derive a wavelength-dependent correction factor by measuring the spectrum of a white light source. White light calibration data was acquired for the all the specimens in the database; however, several of these calibration spectra exhibited unexplained variability. Thus, the spectral response correction was not performed for the database data. This does not present problems for the current analysis because all database data and spectral models were acquired with the same Raman system. However the database data cannot be compared to data acquired using a different Raman system without further calibration steps.

Wavenumber calibration

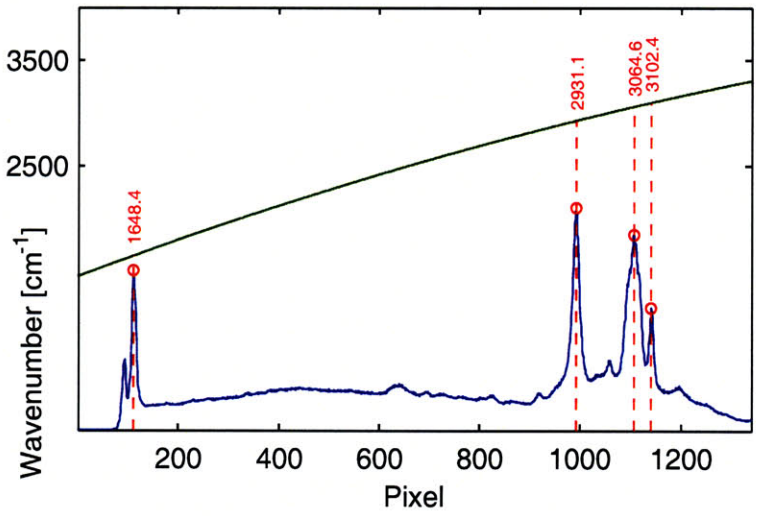
Calibration of the wavenumber axis was performed by mapping CCD pixel number to the vibrational frequencies of a Raman standard. Acetaminophen was chosen as the Raman wavenumber calibration standard due to the large number of vibrational bands covering the fingerprint region [72]. Seventeen Raman bands were used for FP region, and a third order polynomial was used to construct the calibration curve (Figure 2-7(a)). In the HW region, four Raman bands were used, with a second order polynomial (Figure 2-7(b)). Wavenumber calibration was performed independently for each specimen.

Background removal

The acquired spectra are composed of a small Raman signal on top of a large fluorescent background (Figure 2-8), which is likely due to fluorescence emission from structural proteins in the tissue. Background removal in the FP region is challenging because the background varies not only between different specimens but also within a single specimen. One common approach to background subtraction is to estimate the background by fitting the spectral data to a polynomial, and then subtracting the estimated background [83], but this approach was not sufficient for our spectral data. For example, spectra taken from five different sites of a single specimen are shown in Figure 2-8. These spectra illustrate changes in both the fluorescence line shape and in the ratio of Raman-to-fluorescence. Visual inspection of the spectra in Figure 2-8(a) shows that simple polynomial subtraction is not a robust method for extracting the Raman spectrum. The relatively high Raman-to-background ratios in spectra 14 and 23 make it difficult to use a polynomial to estimate the



(a) FP wavenumber calibration



(b) HW wavenumber calibration

Figure 2-7: Wavenumber calibration is accomplished by fitting selected peaks of acetaminophen spectra.

background, because the prominent vibrational peaks in the stronger Raman signal erroneously pull the polynomial fit upward. Furthermore, a single fifth or sixth order polynomial cannot adequately represent the different background signal shapes seen within the specimen. While not all specimens exhibit this much intra-specimen variability, the background spectra between different specimens do exhibit more variability than this.

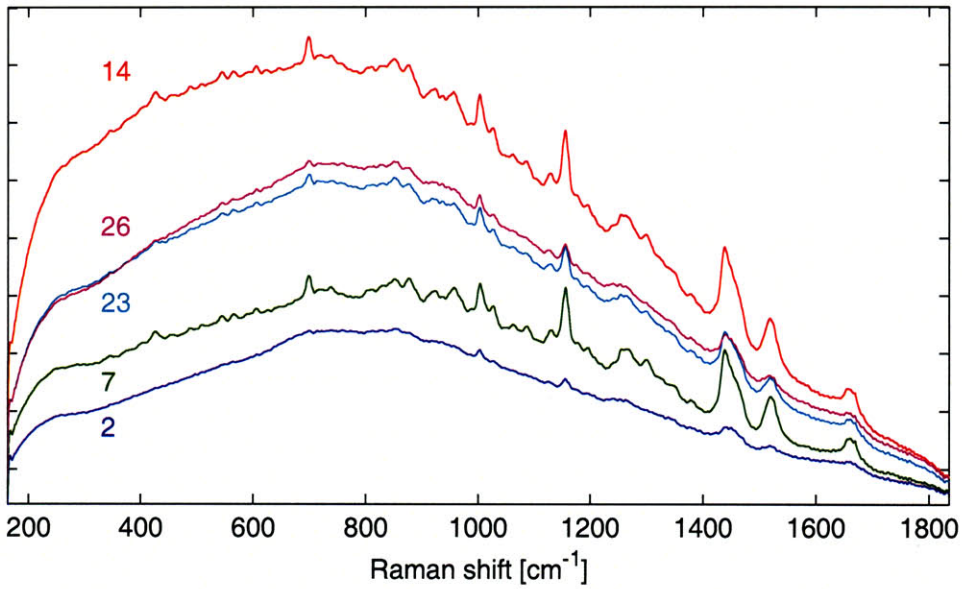
We have heuristically developed a background subtraction routine that addresses the observed variability in both the Raman-to-background ratio and the background line shapes. First, we implemented a modified polynomial fitting routine that iteratively fits a polynomial to the data while excluding sharp Raman peaks [103]. Second, instead of using a single polynomial to fit the entire background spectrum, we fit the background to two distinct polynomials and then performed a spline interpolation to combine them into a smoothly varying background.

In the FP region (Figure 2-9(a)), the background was estimated by a three step process: 1) modified polynomial fits were generated over two distinct regions, from pixel number 240 to 800 with sixth order, and from pixel number 600 to 1340 with fifth order; 2) the polynomials were combined and averaged together in the overlap region from pixel numbers 650 to 750; and 3) the intermediate background was sampled at every 20 pixels and spline interpolated to create a smoothly varying background spectrum. In the HW region (Figure 2-9(b)), the background spectrum was estimated by simply fitting a third order polynomial to the data over a disjoint region encompassing pixels 700 to 890 and 1100 to 1250.

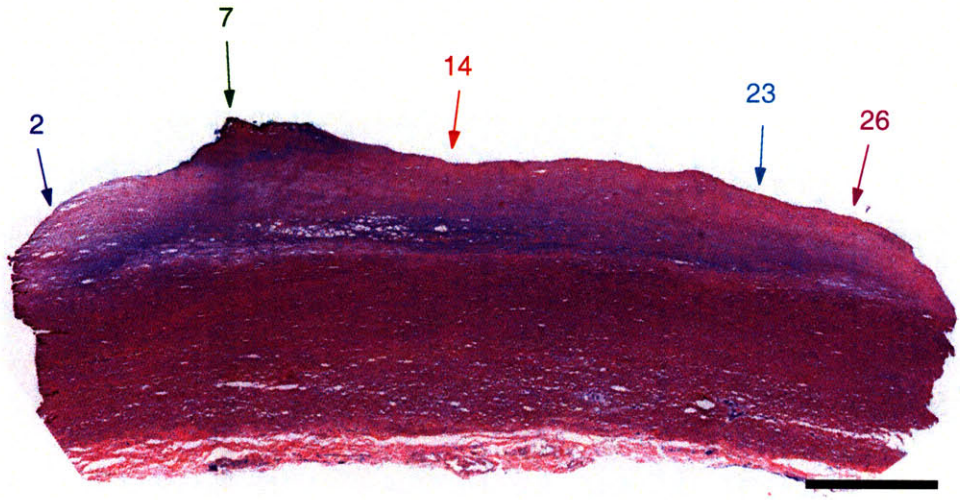
2.2.2 Spectra-histology registration

Tissue spectra identification

The first step in registration is to determine which spectral sites in the Raman line image correspond to the tissue, as opposed to the aluminum specimen holder. This was accomplished by visually inspecting both the preprocessed data and the extracted Raman spectrum. The presence of water in the spectrum as well as the absence or weakness of tissue bands in the Raman spectrum indicates that a spectrum was taken over the aluminum bar. In this way, the site numbers corresponding to the first and last tissue sites were identified.

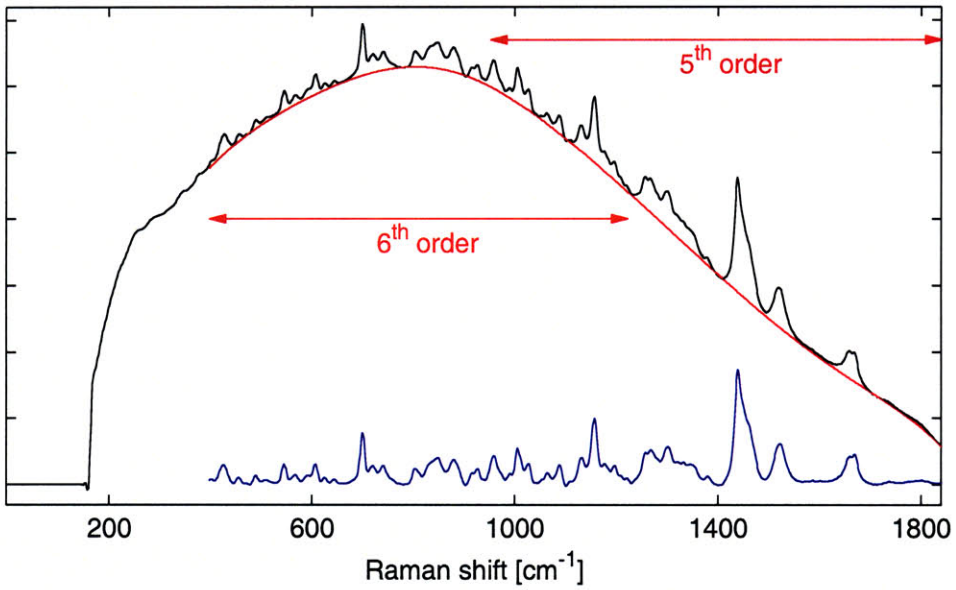


(a) FP Raman spectra at different positions within the same specimen

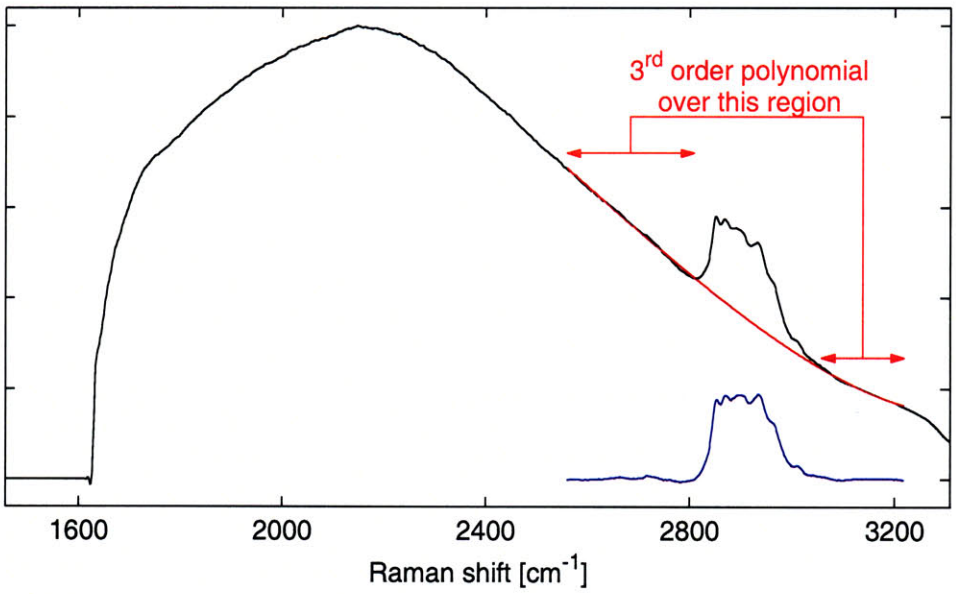


(b) Corresponding histology

Figure 2-8: The preprocessed Raman spectra may have different background shapes, even within the same specimen (UMD089). Scale bar, 1 mm.



(a) FP background removal



(b) FP background removal

Figure 2-9: Raman spectra were extracted from the preprocessed data using a heuristic background subtraction method.

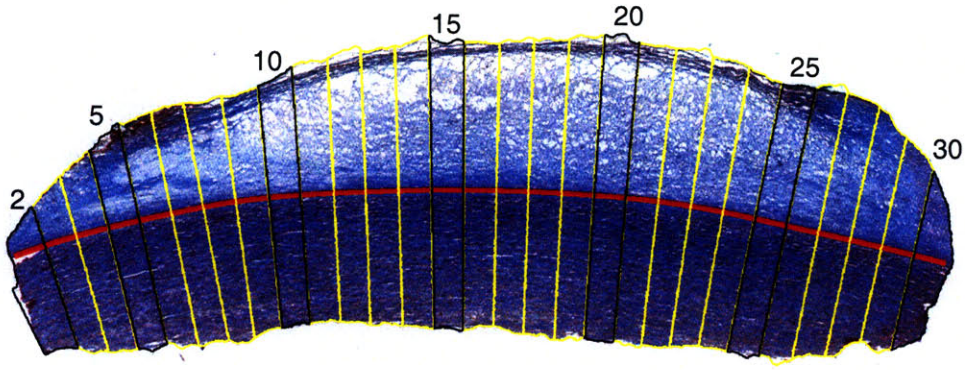


Figure 2-10: Spectra-histology registration. Each vertical region corresponds to a Raman spectrum acquisition site. Black numbers correlate the histology section to the spectrum number. The vertical lines are not intended to imply the actual sampling volume of the spectral measurements.

Histology image preprocessing

For each specimen, the digitized slide was opened in ImageScope, the slide viewing software companion to the Aperio slide scanning system. A low resolution TIFF image was saved, and the image was imported into MATLAB, where it was oriented such that the red edge was on the left of the image and the arterial lumen was at the top of the image (Figure 2-10).

Spectra-histology registration

The final registration step is to assign each Raman tissue spectrum to its spectral acquisition site on the histology image. Because histological processing of thin sections can introduce artifacts such as non-uniform shrinkage and curling, it was not sufficient to simply divide the image into vertical bins. We have developed a registration process that accounts for these histology artifacts. Fore each specimen, a curved line was manually drawn at the boundary between the intima and media, to represent the curvature introduced by histological processing (red line in Figure 2-10). The histology image was then segmented into N equal parts, where N is the number of tissue spectra identified for that specimen. Radial lines were drawn perpendicular to the curved line to denote the individual spectral sites where Raman spectra have been acquired.

2.3 Verification and validation

Before entering a specimen into the *ex vivo* database, the reliability of the spectral and histological data was assessed through several measures. The rigorously standardized procedures and extensive documentation performed throughout data acquisition and spectral and histological data processing enabled the development of an objective verification and validation procedure. Data from each specimen was assessed in the following areas:

- *Spectral quality.* A specimen was rejected if the spectral data exhibited extremely poor SNR, a high fluorescent background that saturated the CCD, or stray light. Additionally, the experimental notes were consulted for anomalies during data acquisition.

- *Registration.* Since accurate registration of the spectral data with the histology sections is a key focus of this database, specimens were rejected if the ink registration marks could not be identified throughout all the histology processing steps. The gross pathology photograph was examined to verify the presence of the registration marks on the tissue. The paraffin embedded histology block was inspected for the presence of green ink, ensuring that the thin sections were cut from the correct surface of the specimen. The stained thin sections were also inspected for the presence of both green and red ink. Additionally, the gross pathology photo was compared to the histology block and the thin sections to ensure that specimens were not misidentified. Again, the experimental notes were consulted to identify anomalies encountered during the acquisition procedure. Specimens that did not meet all of these requirements were rejected from the database.

- *Histology quality.* Specimens were rejected if the histological sections were uninterpretable, either due to poor sectioning or poor staining. Histology quality was assessed by a pathologist who was blinded to the spectroscopic data.

These verification measures ensured that all of the specimens in the *ex vivo* Raman database have reliable data.

2.4 Pathological diagnosis

A board-certified pathologist diagnosed each specimen using the trichrome and H&E stained slides. A modified version of the classification scheme proposed by Virmani *et al.* was

used [11] (Table 2.1). Some specimens displayed heterogeneous lesions; in these cases the slide was classified with the most clinically significant category. In some cases, additional stains were requested and supplied to confirm a diagnosis, such as picrosirius red, which can be used to quantify collagen content, and Movat's pentachrome, to establish the presence of proteoglycans.

Table 2.1: Histological classification scheme used for the database. This scheme is adapted from [11].

Category		Description
Intimal hyperplasia	IH	Thickening of the intimal region. Predominantly collagen and smooth muscle cells with $< 10\%$ lipid or foam cells present. Intimal thickness should be less than $\sim$300 μm to be classifed as IH.
Fibrous plaque	FP	Fibrous lesions consisting predominantly of collagen and smooth muscle cells. These lesions have negligible or no lipid ($< 40\%$ lipid within the ROI). The thickening should be greater than $\sim$300 μm to be considered FP.
Pathological intimal thickening	PIT	Lipid rich plaque with no visible necrosis.
Necrotic core fibroatheroma	NCFA	Necrotic lipid rich core with overlying fibrous cap.
Calcified	FC	Lesions with calcified nodules.

In addition to providing a diagnosis, the pathologist was asked to identify histological features such as necrotic core, calcification, and protein-rich areas. Using ImageScope's annotation feature, morphological features were directly identified on the image (Figure 2-11). Additional comments from the pathologist were noted within the software. ImageScope also provides a calibrated scale for measuring the size of morphological features, such as the thickness of a fibrous cap. All information was saved in an XML file, which was used for later processing in MATLAB.

Digital capture of the histology image and the pathologist's diagnostic notes is a key feature of the *ex vivo* Raman database, facilitating mining of the database.

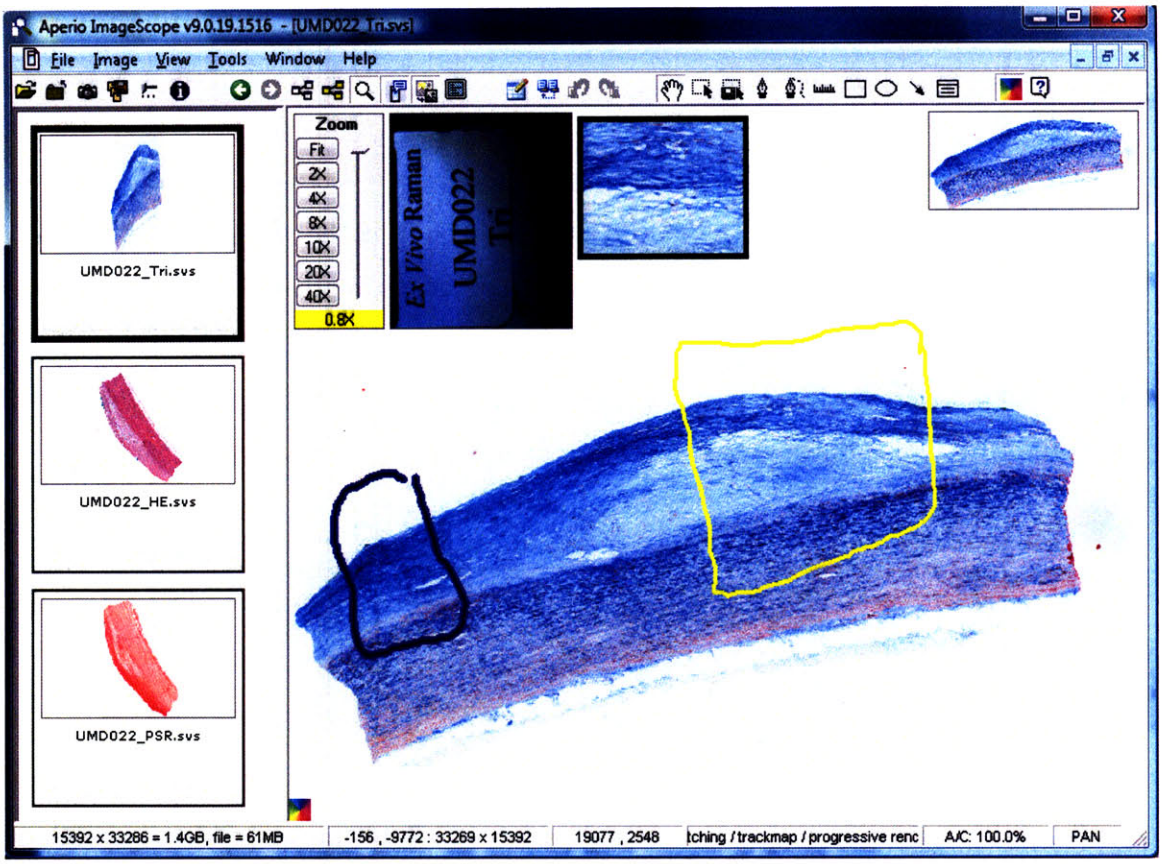


Figure 2-11: Using Aperio's ImageScope software, a pathologist has identified distinct regions of a heterogeneous plaque (UMD022). The yellow box encloses a necrotic core region and the blue box encloses a protein-rich region.

2.5 Creation of the *ex vivo* Raman database

The database was created using the methods described above. For each cadaver aorta that was procured, several atherosclerotic lesion specimens were scanned with the Raman line imaging system. The protocol for each specimen is summarized as follows:

1. Procure and prepare lesion specimen; document with notes.
2. Acquire calibration data.
3. Acquire Raman line image.
4. Prepare specimen histology.
5. Extract Raman spectra from raw data.
6. Register spectra with histology slide.

7. Validate data quality.

8. Assign pathological diagnosis.

To date, ninety specimens from ten cadaver aortas have been scanned. Twenty-one specimens were rejected from inclusion in the database during the validation step. Nine of the rejected specimens were the first nine scanned specimens, representing a learning curve and the need for minor adjustments to the protocol. A group of specimens from subject 8 were rejected, mainly due to quality concerns in the histology sections returned from MGH Pathology. The remaining 69 samples were diagnosed by the pathologist (Table 2.2).

Table 2.2: *Ex vivo* Raman database disease distribution. Number of lesions.

Subject	IH	FP	PIT	NC	CA	Total
1	1	3	8	4	0	16
2	0	2	5	0	0	7
3	0	0	2	0	0	2
4	0	0	3	0	0	3
5	1	0	3	0	1	5
6	0	0	2	1	1	4
7	0	0	8	2	1	11
8	0	1	0	0	0	1
9	0	0	18	2	0	20
Total	2	6	49	9	3	69

The database includes 69 Raman line images from nine subjects. Because one of the initial goals of this thesis is to more thoroughly characterize the broad class of "lipid-rich" lesions, the database was deliberately skewed to include a large proportion of lipid-rich specimens. Thus, lipid-rich lesions constitute 85% of the database specimens, while protein-rich (collagen) and calcified lesions represent only 12% and 4%, respectively.

2.6 Visualization

The *ex vivo* database incorporates data from several different sources, including spectral data in two wavenumber regions, histology sections, and histopathological diagnoses and notes. We have developed a graphical user interface (GUI) that enables the user to visualize the data in the database (Figure 2-12). For each spectral site, the GUI displays the acquired raw spectra, the extracted FP and HW Raman spectra, and the corresponding site outlined

in the histology section. The spectral filenames are displayed in the upper box, and the pathology diagnosis and any relevant notes are shown to the left of the histology image. Check boxes allow the user to view the raw and processed spectra on a fixed intensity scale and to view spectral model fits, residuals, and fit coefficients (to be described in Chapter 3). Using the "Back" and "Next" buttons in the upper right, the user can quickly navigate through all the acquisition sites in the specimen.

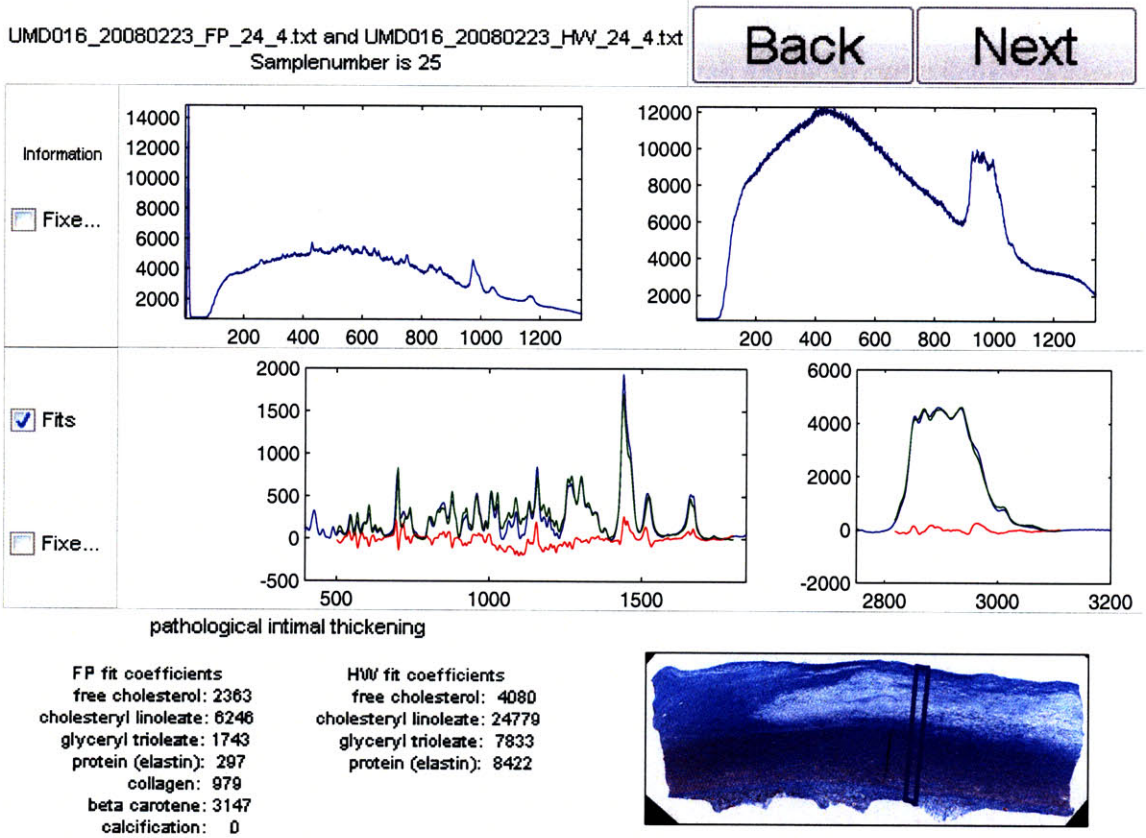


Figure 2-12: GUI for viewing the database.

The GUI has been an invaluable tool for managing and analyzing the data in the *ex vivo* Raman database. The ability to quickly visualize all the data from an entire specimen on a single screen has aided in correlating the spectral information obtained from the Raman line images with the morphological information obtained from the histology images.

2.7 Conclusions

In this chapter, we have developed an *ex vivo* Raman database consisting of Raman line image data and histological data from 69 aortic lesions. Raman spectra were acquired at fixed intervals across each lesion, at physiological temperature, in both the fingerprint and high wavenumber regions. A rigorous protocol was developed to ensure that all specimens had good quality spectral data, good quality histological data, and most importantly, accurate registration between the spectral acquisition sites and the histology image. The histology sections and the pathologist's diagnosis and notes were digitally captured. All the spectral and histological data was stored electronically, and a graphical user interface was developed to enable visualization of all the data for a specimen on a single screen. In the next chapter, we focus on the lesion characterization using the database data.

Chapter 3

Ex Vivo Lesion Characterization

Having created the *ex vivo* Raman database of correlated spectra and histology described in Chapter 2, we turn our attention to characterizing atherosclerotic plaques based on the Raman line images. In this chapter, we develop a set of physically motivated basis spectra to aid in interpreting the tissue spectra. Each spectrum in the line image is fit to the spectral model, and the resulting fit coefficients are used to diagnose lesions based on single site Raman spectra and to correlate the spectral information in Raman line images with the histological sections.

3.1 Spectral modeling

As discussed in Section 1.2, an important step in Raman spectral analysis is data reduction, the process of reducing the multi-dimensional Raman spectrum into a smaller set of relevant parameters. In this chapter we model each tissue Raman spectrum as a linear combination of physical basis spectra, an approach that has been successfully demonstrated for characterizing atherosclerotic plaques [67, 68, 70, 83, 91].

Baraga and Manoharan *et al.* determined that for arterial Raman spectra, most spectral features can be adequately represented as a linear combination of the following chemical components: collagen, elastin, cholesterol, cholesteryl oleate, cholesteryl linoleate, and calcium hydroxyapatite [67, 104]. They created a basis spectra model using commercially available preparations of these chemicals, and then experimentally determined the relative Raman cross section of each chemical, resulting in a calibrated model capable of quantitatively determining the chemical composition of a sample based on its Raman spectrum.

They experimentally validated the calibrated model using mixtures of collagen, elastin, cholesterol, cholesteryl oleate, and cholesteryl linoleate and were able to use Raman spectroscopy to determine the composition to within $5-15\%$ of the actual value.

Noting that the Raman activity of a commercially prepared chemical may not accurately reflect the Raman activity of the chemical *in situ*, Brennan *et al.* refined the basis spectra model, using basis spectra of chemical compounds extracted from human coronary artery [68]. This model required seven components: three lipids, two proteins, calcium salts, and β-carotene. The three lipids were free cholesterol, cholesterol ester, and triglycerides; the cholesterol spectra were obtained from purified chemicals, while the triglyceride spectrum was measured after chemically extracting triglycerides from adventitial adipose tissue. For the two protein spectra, noncalcified arteries were chemically delipidized and Raman spectra were acquired. After examining several arteries, they isolated two representative protein spectra that modeled all the spectra well. The calcium salt spectrum was obtained by dissecting a calcified nodule from a diseased artery, and the β-carotene spectrum was measured from crystalline β-carotene. They applied this model to minced coronary artery samples, using biochemical assays to determine the absolute concentration of each chemical component. The Raman spectral model was capable of assessing chemical composition to within a few percentage points.

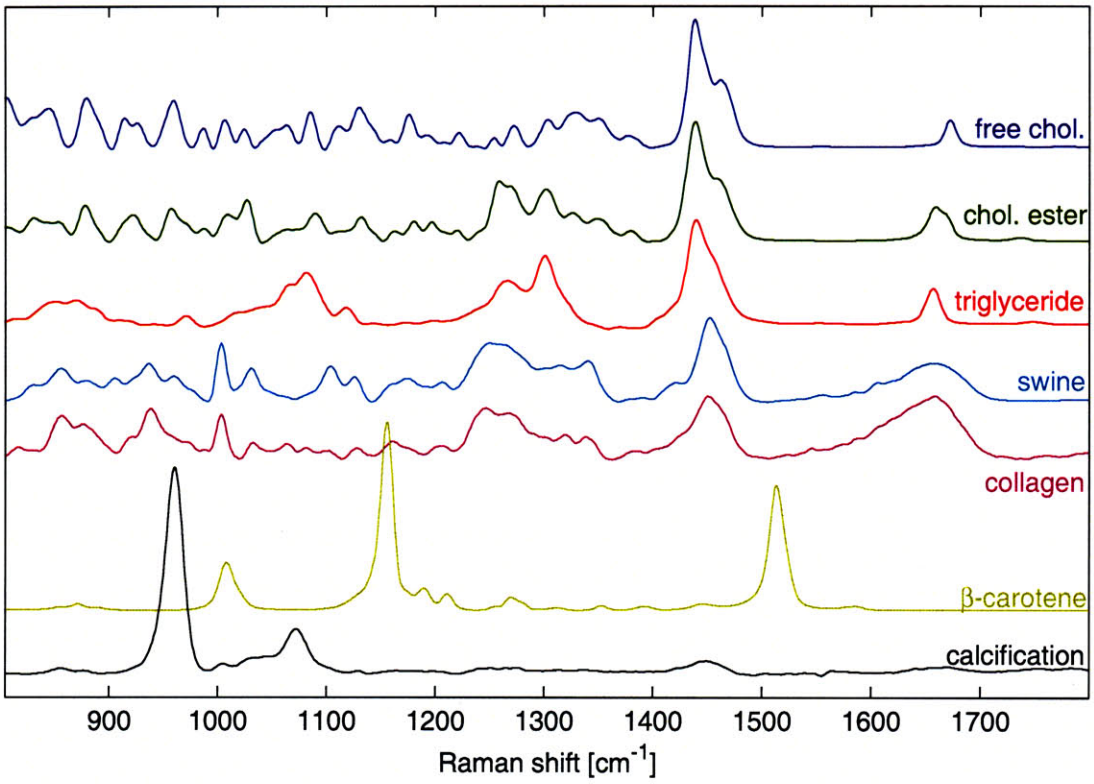
Buschman *et al.* used a confocal Raman system to acquire basis spectra from the most important morphological components of arterial tissue *in situ* [91]. From thin sections of coronary arteries, they isolated and acquired Raman spectra for the following nine components: internal and external elastic lamina, collagen fibers, smooth muscle cells, adventitial fat, foam cells, necrotic core, cholesterol crystals, carotene containing crystal, and calcification. The spectra for these morphological components were fit to a model consisting of purified chemical spectra, similar to the models used by Baraga and Manoharan [67, 104]. In this analysis, they did not use a calibrated model, and thus did not assess quantitative chemical composition, only relative fit contribution to the Raman spectrum. They found that elastic lamina are primarily composed of elastin and collagen (~2:1 ratio); collagen fibers are primarily collagen; smooth muscle cells are primarily actin; and adventitial fat is primarily triolein. Foam cells and necrotic core are very similar and are composed of cholesterol linoleate, cholesterol, and collagen (~2:1:1 ratio); cholesterol crystals are primarily free cholesterol and cholesterol linoleate (~2.5:1 ratio); carotene containing crystals

are overwhelmingly β-carotene; and calcifications are primarily composed of calcium hydroxyapatite. Additional chemicals were included in the biochemical model, including myosin, tropomyosin, phosphatidyl choline, and calcium carbonate, but these contributed little to the morphological spectra.

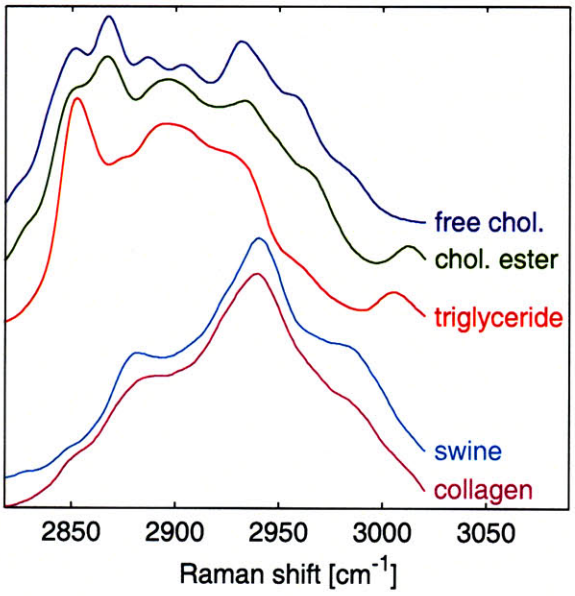
3.1.1 Physical basis spectra model

Based on the results of the previously developed models, we have created a spectral model containing the following components: free cholesterol, cholesterol ester, triglyceride, proteins, β-carotene, and calcification. Where possible, we obtained basis spectra *in situ* (proteins and calcification), but for the remaining components, commercially purified chemicals were used. All spectra were acquired using the benchtop system described in Chapter 2. After acquisition, the basis spectra were preprocessed as described in Section 2.2.1 and background removal was performed analogously to Section 2.2.1. The background-removed basis spectra are shown in Figure 3-1. While the spectrometer acquires data in the wavenumber range from 300 to 1840 cm^{-1} in the FP region and 1600 to 3300 cm^{-1} in the HW region, the basis spectra have been truncated to a smaller region. In the FP region, the spectra are truncated to $804 - 1799$ cm^{-1} because the background removal procedure is unreliable outside this region. In the HW region, we consider only the region between 2818 and 3021 cm^{-1}, because this region exhibits high variability between the basis spectra. Beyond this region, the HW basis spectra exhibit flat tails.

Cholesterol is a specific molecule, so purified cholesterol (Sigma C8667, Sigma-Aldrich; St. Louis, MO) was used to measure the free cholesterol Raman spectrum. The term "cholesterol esters" describes cholesterol esterified to any one of several fatty acids. In our spectral model, the class of cholesterol esters is represented by cholesteryl linoleate (Sigma C0289). Triglycerides are represented in our model by glyceryl trioleate (Sigma T7140). The lipid spectra are very similar to previously published spectra [71]. Both cholesterol and cholesteryl linoleate exhibit a strong band at 1439 cm^{-1} due to CH_2 bending. The C=C stretching vibration of the sterol ring results in a band at 1671 cm^{-1} in cholesterol, which is shifted to 1659 cm^{-1} in the cholesteryl linoleate spectrum. Several bands arise from vibrations in the sterol nucleus, including the cholesterol bands at 958 and 873 cm^{-1} and the cholesteryl linoleate bands at 957, 922, and 878 cm^{-1}. The glyceryl trioleate spectrum also corresponds very well with published spectra [71], with prominent peaks at 1440 cm^{-1}



(a) Fingerprint basis spectra



(b) High wavenumber basis spectra

Figure 3-1: Chemical basis spectra in the (a) fingerprint and (b) high-wavenumber regions. Only the displayed wavenumber regions (FP: $804 - 1799$ cm^{-1}, HW: $2818 - 3021$ cm^{-1}) were used in the basis spectra model.

due to CH_2 bending, at 1656 cm^{-1} due to C=C stretching, and at 1301 and 1265 cm^{-1} due to =C-H bending.

The most important structural proteins to include in the spectral model are elastin and collagen, but it is difficult to acquire Raman spectra of commercially purified elastin and collagen due to the presence of a very large fluorescent background. Thus, we represent both elastin and collagen using two protein spectra. First, we acquired spectra from normal swine aorta *ex vivo*.[1] Biochemically, swine aorta is primarily elastin and collagen, with the ratio of elastin to collagen decreasing with increasing distance from the heart [105]. The swine aorta used in our model was harvested close to the swine's heart, so we believe that our swine spectrum is predominantly composed of elastin with an additional collagen component, though we do not know in what proportion. The second protein spectrum included in our model was acquired from a tissue-engineered cartilage sample[2] and is predominantly collagen [106].

Raman spectra of the two proteins are similar, with both exhibiting prominent bands around 1450 cm^{-1} due to the CH_2 bending mode of proteins, at 1664 cm^{-1} due to the amide I band, at 1264 cm^{-1} due to the amide III band [107, 108], and at 1004 cm^{-1} due to phenylalanine [91]. Comparing the swine spectrum to that of internal elastic lamina published by Buschman *et al.* [91], the swine spectrum has bands at 1136 and 1104 cm^{-1} due to desmosine/isodesmosine, which is specific to elastin [109]. The swine spectrum also has bands at 933 and 855 cm^{-1}, due to the C-C stretching mode of proline; these bands are specific to collagen [91, 109]. Thus, as expected, the swine spectrum is a combination of elastin and collagen. The collagen spectrum is missing the elastin-specific band at 1104 cm^{-1}, and the 855 cm^{-1} band is more pronounced relative to the swine spectrum, confirming that the cartilage spectrum approximates pure collagen.

To obtain a spectrum representing atherosclerotic calcification *in situ*, we dissected a large, exposed calcified nodule from an *ex vivo* human aorta specimen. The calcified spectrum is dominated by a band at 960 cm^{-1}, due to the symmetric PO_4^{3-} stretch of calcium hydroxyapatite. Another band at 1071 cm^{-1} is attributed to the symmetric carbonate (CO_3^{2-}).

[1] Normal swine aorta was harvested from a swine sacrificed for another procedure, under approval from the Massachusetts General Hospital's Subcommittee on Research Animal Care (protocol #2007N000041).

[2] The tissue-engineered cartilage sample was a generous loan from the Plastic Surgery Research Laboratory at Massachusetts General Hospital.

The β-carotene basis spectrum was acquired from crystallized β-carotene (Sigma Fluka BioChemika 22040). The β-carotene basis spectrum exhibits three resonantly enhanced bands at 1513 cm^{-1} due to the C=C stretching mode, at 1157 cm^{-1} due to the C-C stretching mode, and at 1008 cm^{-1} due to the C-H bending mode [110]. β-carotene *in situ* is typically associated with lipids, which shifts the C=C stretching band to ~1524 cm^{-1}. This highlights one of the challenges in developing a physically motivated spectral model: purified chemicals may behave differently in an artery *in situ*, leading to shifts, linewidth broadenings, or intensity variations of the Raman bands.

Raman spectra in the high wavenumber region correspond well with published spectra [88]. The Raman bands between 2800 and 3050 cm^{-1} are attributed to CH stretching modes. The free cholesterol and cholesterol ester spectra both have strong bands at 2867 cm^{-1}, while the triglyceride spectrum has a corresponding band at 2852 cm^{-1}. In contrast, the proteins exhibit a strong band ~2940 $cm^{-1}$and a shoulder near 2990 cm^{-1}. Neither the calcified nodule nor the β-carotene contained any prominent bands in the high wavenumber region, so they are not included in the HW spectral model.

Our spectral model is missing a number of chemicals that are known to be present in the arterial wall, including actin, phospholipids, and other cholesterol esters, particularly cholesteryl oleate. While the absence of these components in the basis spectra model may affect the accuracy of the decomposition, these components are not believed to contribute much to the overall Raman signal from atherosclerotic plaque [91].

Orthogonality of the basis spectra

The accuracy with which we can fit an arbitrary arterial Raman spectrum to the basis spectra model is determined in part by the degree of orthogonality between the bases. Orthogonality between two vectors is assessed by calculating the correlation coefficient,

$$\rho = \frac{\boldsymbol{x} \cdot \boldsymbol{y}}{\|\boldsymbol{x}\| \, \|\boldsymbol{y}\|}, \tag{3.1}$$

where $\boldsymbol{x}$ and $\boldsymbol{y}$ are the two vectors of interest. We assess the orthogonality of the basis spectra by calculating the correlation coefficient between each pair of chemical basis spectra (Table 3.1).

In the FP region, free cholesterol and cholesterol ester are highly correlated with $\rho =$

Table 3.1: Orthogonality of the basis spectra

(a) Correlation coefficients for the FP basis spectra (Figure 3-1(a)).

		Lipids			Proteins			
		free chol.	chol. ester	triglyceride	elastin	collagen	β-carotene	calcification
Lipids	free chol.	1.00	0.81	0.65	0.45	0.29	-0.08	0.16
	chol. ester	0.81	1.00	0.85	0.68	0.49	-0.10	0.11
	triglyceride	0.65	0.85	1.00	0.59	0.45	-0.13	0.03
Proteins	elastin	0.45	0.68	0.59	1.00	0.82	-0.13	0.06
	collagen	0.29	0.49	0.45	0.82	1.00	-0.17	0.08
	β-carotene	-0.08	-0.10	-0.13	-0.13	-0.17	1.00	-0.10
	calcification	0.16	0.11	0.03	0.06	0.08	-0.10	1.00

(b) Correlation coefficients for the HW basis spectra (Figure 3-1(b)).

		Lipids			Proteins	
		free chol.	chol. ester	triglyceride	elastin	collagen
Lipids	free chol.	1.00	0.97	0.89	0.39	0.53
	chol. ester	0.97	1.00	0.95	0.33	0.48
	triglyceride	0.89	0.95	1.00	0.25	0.40
Proteins	elastin	0.39	0.33	0.25	1.00	0.98
	collagen	0.53	0.48	0.40	0.98	1.00

0.81, and triglyceride is also highly correlated with cholesterol ester, with $\rho = 0.85$. The two proteins are highly correlated with a correlation coefficient of $\rho = 0.82$. As a group, lipids are moderately correlated with each other, but not correlated with the proteins. The final two basis spectra components, β-carotene and calcification, exhibit very low correlation with all other components. Thus, we expect the FP basis spectra model to be capable of accurately distinguishing between lipids, proteins, β-carotene and calcification, but it may not be as accurate for distinguishing between the two proteins, or between the three lipids.

Not surprisingly, a similar pattern is seen in the HW basis spectra. The two protein components are highly correlated ($\rho = 0.98$), which is is easily confirmed by observing the striking visual similarities between the swine and collagen spectra in Figure 3-1(b). Additionally, the free cholesterol and cholesterol ester basis spectra are highly correlated as well ($\rho = 0.97$), which can also be visually observed (Figure 3-1(a)). However, the highest correlation between the lipids and proteins in the HW region is relatively low at $\rho = 0.53$, suggesting that the HW basis model is also suitable for distinguishing between the lipid and protein components of an arterial Raman spectrum.

3.1.2 Spectral decomposition

We model each tissue spectrum, $S(\tilde{\nu})$, as a linear combination of the physical basis spectra, *i.e.*,

$$\hat{S}(\tilde{\nu}) = \sum_{k \in \mathbf{B}} \alpha_k S_k(\tilde{\nu}) \tag{3.2}$$

where $\hat{S}(\tilde{\nu})$ is the reconstructed tissue spectrum, $S_k(\tilde{\nu})$ is the basis spectrum for component k, α_k is the weighting or fit coefficient for basis component k, and $\mathbf{B}$ is the set of basis spectra. In the fingerprint region, $\mathbf{B_{FP}}$ = {free chol., chol. ester, triglyceride, swine, collagen, β-carotene, calcification}, and in the high wavenumber region, $\mathbf{B_{HW}}$ = {free chol., chol. ester, triglyceride, swine, collagen}. Prior to spectral decomposition, each basis spectrum was normalized to have unit energy over the wavelength region shown in Figure 3-1. We define the residual as the difference between the tissue spectrum and the reconstructed spectrum,

$$R(\tilde{\nu}) = S(\tilde{\nu}) - \hat{S}(\tilde{\nu}). \tag{3.3}$$

Each tissue spectrum, $S(\tilde{\nu})$, was preprocessed and had its background removed, but no filtering was performed prior to decomposition. We solve Eq. (3.2) as a constrained linear

least-squares minimization problem, with the constraint that the weighting coefficients, α_k, must all be positive. The fit coefficients were found using a nonnegative least-squares routine in MATLAB (`lsqnonneg`).

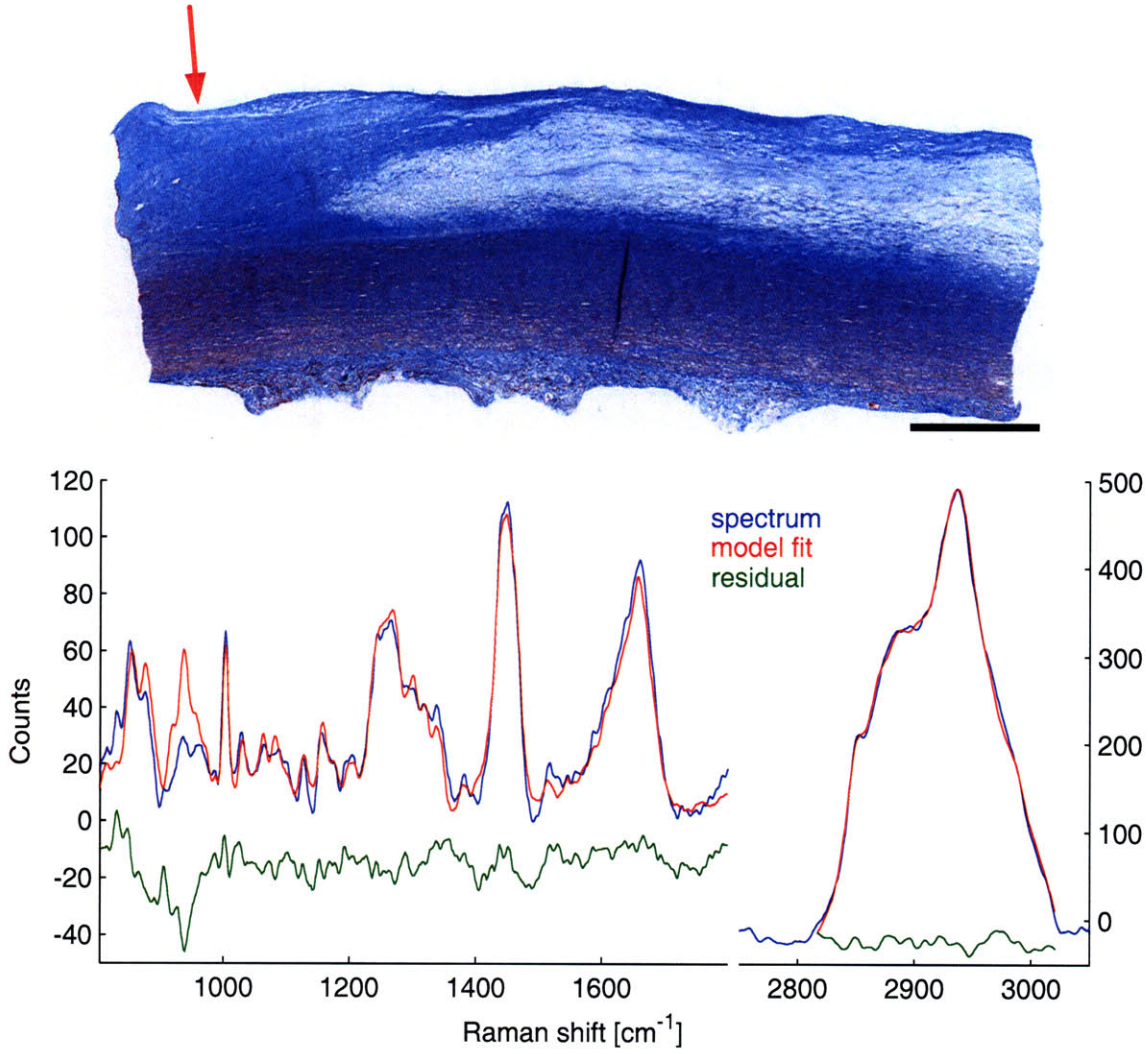
Example spectra are shown in Figures 3-2, 3-3, and 3-4. The spectra in these figures have been filtered for display, and the residuals have been offset for clarity. Fit coefficients and normalized fit coefficients are shown for each sample. Since we have not attempted to calibrate the spectral model to account for varying Raman cross sections, the normalized fit coefficients do not indicate the physiological concentration of each basis component.

An example of a protein-rich spectrum is shown in Figure 3-2. This heterogeneous specimen (UMD016) was diagnosed with pathological intimal thickening (PIT), containing a protein-rich region (dark blue stained area) at the left edge of the sample that gradually transitions into a more lipid-rich region (gray stained area). The basis spectra reconstruct the tissue spectrum moderately well for both wavenumber regions. In the FP region, all of the significant bands in the tissue spectra are represented by the bases, but the band intensities do not all match, leading to peaks that can be seen in the residual. In particular, the spectral model overfits the region between 900 and 980 cm^{-1}. This is seen consistently throughout the database and is possibly due to poor background estimation in the protein basis spectra. In the HW region, the basis spectra fit the data well.

The normalized fit coefficients confirm that this spectral site is protein-rich, with high normalized fit coefficients for the protein bases in both the FP and HW regions. However, the swine and collagen fit coefficients are not consistent between the FP and HW regions. The swine fit coefficient is 0.01 in the FP region and 0.1 in the HW region, while the collagen fit coefficient is 0.68 in the FP region and 0.58 in the HW region. The discrepancy highlights the difficulties encountered when attempting to decompose a spectrum into highly correlated basis spectra.

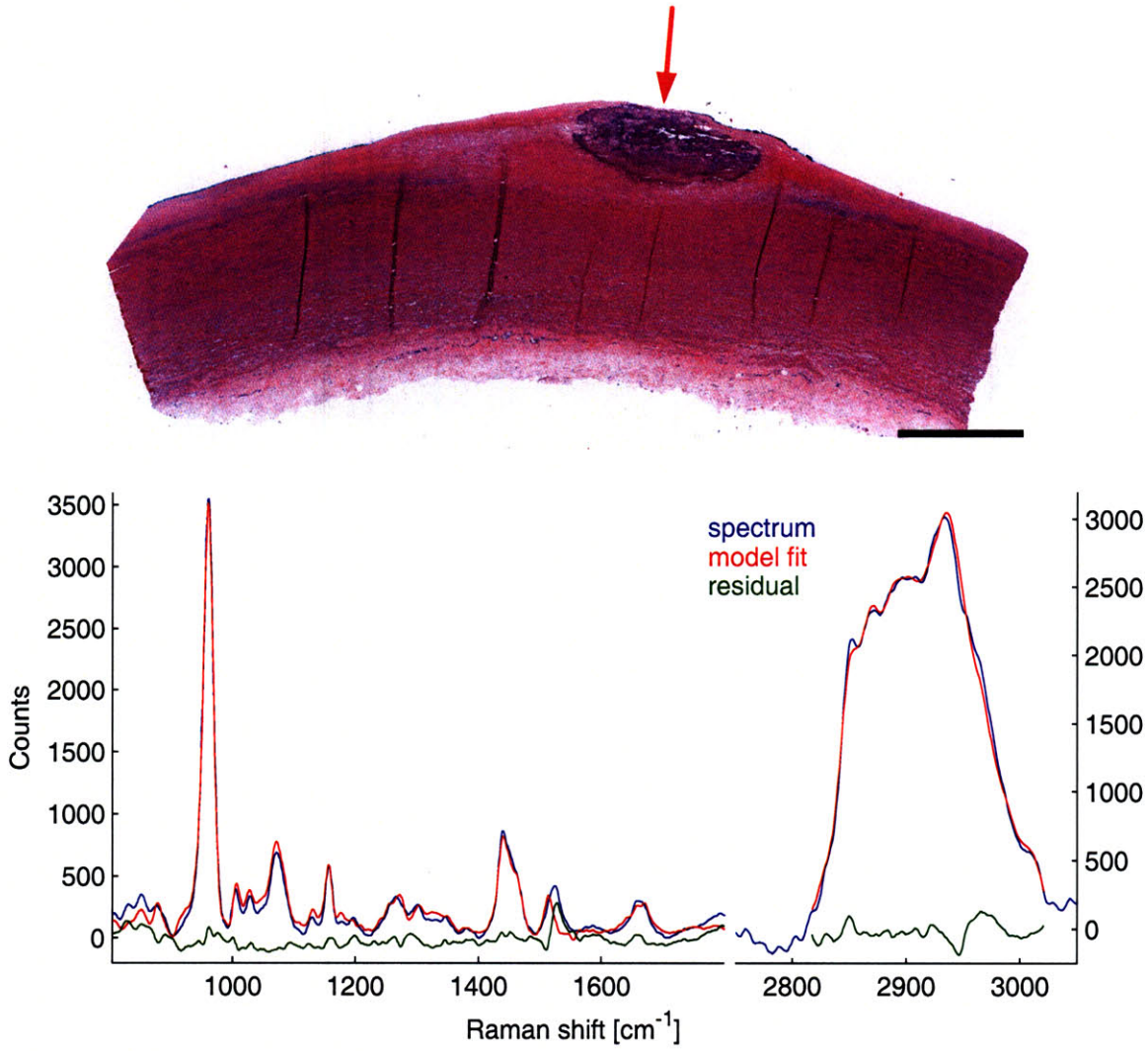
An example of a calcified spectrum (specimen UMD045) is shown in Figure 3-3. The calcified nodule is easily identified as the purple region in the H&E histology section. The reconstructed fit spectrum accurately represents the PO_4^{3-} band at 960 cm^{-1}, which is not surprising because that band is a strong feature that is orthogonal to all the other basis spectra. Since the calcification has no signature in the HW region, the HW fit coefficients represent only the relative contributions of cholesterol, triglyceride, and protein.

Finally, an example of a necrotic core spectrum (specimen UMD087) is shown in Figure



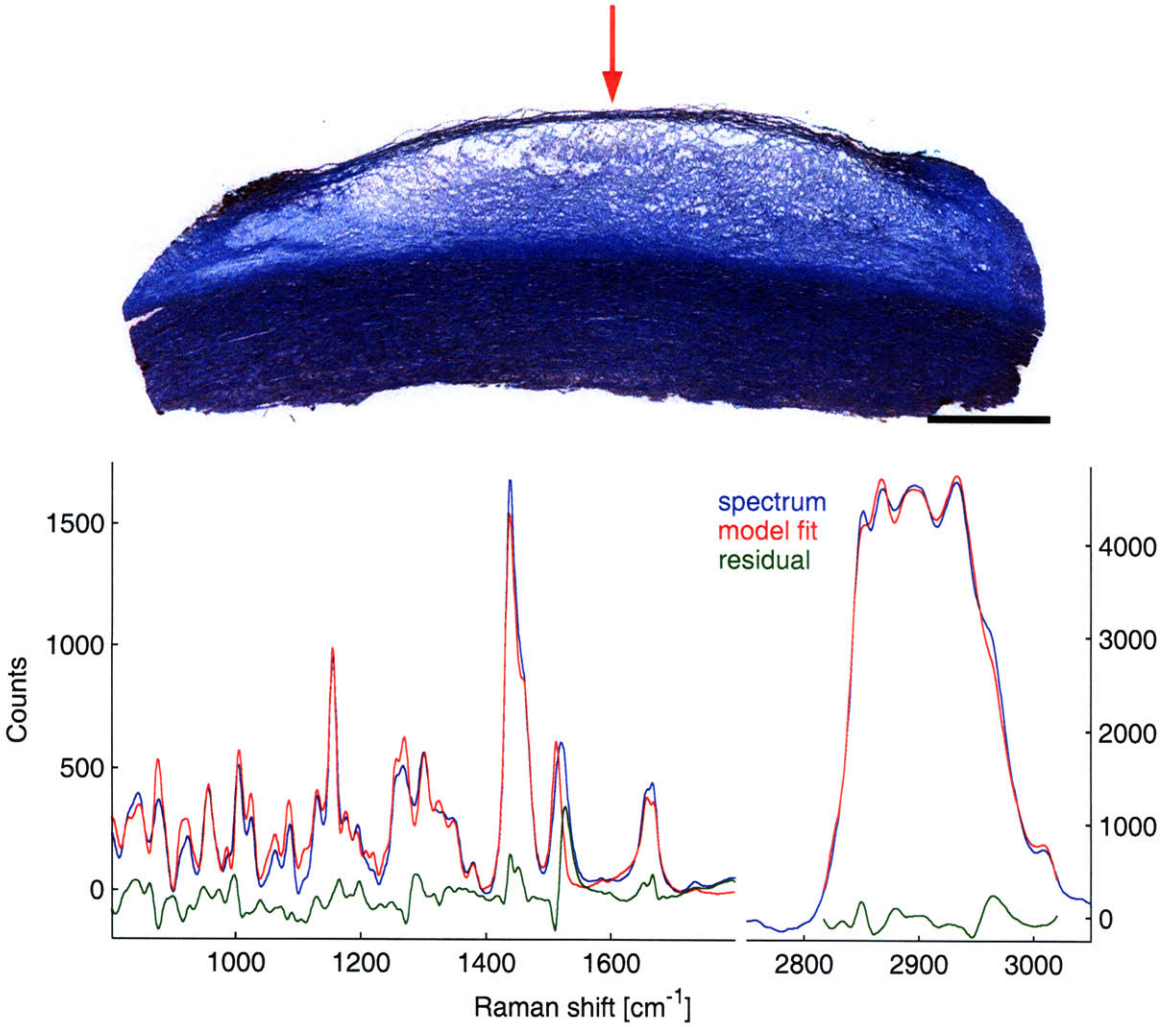
	Fingerprint		High wavenumber	
Chemical	raw	norm.	raw	norm.
free chol.	53	0.04	688	0.17
chol. ester	118	0.10	171	0.04
triglyceride	141	0.12	440	0.11
swine	15	0.01	422	0.10
collagen	810	0.68	2372	0.58
β-carotene	52	0.04	–	–
calcification	0	0.00	–	–

Figure 3-2: Basis spectra fit for a fibrous region of a plaque (UMD016). The red arrow in the trichrome image marks the spectral acquisition site, a protein-rich region (stained blue) of a heterogeneous plaque diagnosed as pathological intimal thickening. The tissue spectrum is shown with the reconstructed fit spectrum and the corresponding residual (offset). Raw and normalized fit coefficients are given in the table. Scale bar, 1 mm.



	Fingerprint		High wavenumber	
Chemical	raw	norm.	raw	norm.
free chol.	1634	0.09	1266	0.05
chol. ester	1607	0.09	9206	0.36
triglyceride	193	0.01	3334	0.13
swine	0	0.00	0	0.00
collagen	846	0.05	11755	0.46
β-carotene	2006	0.11	–	–
calcification	11615	0.65	–	–

Figure 3-3: Basis spectra fit to a calcified region of a plaque (UMD045). A large calcification, stained purple, can be identified in the H&E histology image. The red arrow denotes the spectral acquisition site. The tissue spectrum is shown with the reconstructed fit spectrum and the corresponding residual (offset). Raw and normalized fit coefficients are given in the table. Scale bar, 1 mm.



	Fingerprint		High wavenumber	
Chemical	raw	norm.	raw	norm.
free chol.	3050	0.22	5374	0.12
chol. ester	4716	0.35	23639	0.54
triglyceride	866	0.06	6046	0.14
swine	0	0.00	8451	0.19
collagen	1463	0.11	0	0.00
β-carotene	3567	0.26	–	–
calcification	0	0.00	–	–

Figure 3-4: Basis spectra fit to a necrotic core plaque (UMD087). The red arrow in the trichrome image denotes the spectral acquisition site. The tissue spectrum is shown with the reconstructed fit spectrum and the corresponding residual (offset). Raw and normalized fit coefficients are given in the table. Scale bar, 1 mm.

3-4. There is a strong correspondence between the lipid and protein fit coefficients in the FP and HW regions. Both regions are dominated by free cholesterol and cholesterol ester, with a moderate protein component present. The reconstructed fit spectra represent all the major peaks present in the tissue spectra in both regions.

3.1.3 Discussion

Peak Raman intensity in the high wavenumber region is consistently higher than in the fingerprint region (at the $1440 - 1450$ cm^{-1} CH_2 bend), due to the differing Raman cross sections between the CH_2 bending mode (FP) and the CH stretching modes (HW). Thus the HW Raman spectra typically exhibit higher SNR than the FP spectra.

Additionally, the overall Raman intensity is very low for the protein-rich spectrum relative to the lipid rich spectrum. The protein-rich spectrum has peak intensities of only 112 counts at 1450 cm^{-1} (CH_2) and 491 counts at 2937 cm^{-1}, while the necrotic core spectrum has peak intensities of 1685 at the corresponding 1439 cm^{-1} band, and 4743 at 2934 cm^{-1}. This is consistent with the fact that the Raman cross sections for lipids are known to be higher than the protein cross sections [104]. This suggests that Raman intensity may be a useful metric for lesion characterization; however, Raman spectra acquired through a catheter *in vivo* may be susceptible to intensity variation due solely to geometry. For example, a Raman spectrum acquired when holding a catheter in contact with a lesion would be greater in magnitude than a spectrum acquired with the catheter held 1 mm away from the tissue surface. For this reason, we do not consider the signal intensity in this analysis.[3]

The β-carotene basis spectrum does not accurately represent the tissue spectra in the 1500 to 1540 cm^{-1} region. In both the calcified and necrotic core samples (Figures 3-3 and 3-4), the band at ~1520 cm^{-1} is poorly fit by the 1512 cm^{-1} band in the β-carotene basis spectrum, giving rise to a sharp feature in the residuals. This can be explained by considering that the β-carotene basis spectrum was acquired from a dry, purified powder, which does not accurately represent the chemical state of β-carotene in an artery *in situ*, where it is generally associated with lipid domains, due to its highly lipophilic nature. Fortunately, no other basis spectra have spectral features in the 1500 to 1540 cm^{-1} region, so the lack of an accurate basis in this region does not affect the other fit coefficients.

[3]It is possible to devise a clinical procedure that *would* yield reliable Raman spectra intensity information; in that case, absolute Raman intensity is likely to be an important parameter for lesion characterization.

Furthermore, the other prominent feature of the β-carotene spectrum, the band at 1156 cm^{-1}, does fit the tissue spectra well, suggesting that the β-carotene fit coefficients are trustworthy.

Overall, the FP and HW basis spectra models do not fit the tissue spectra perfectly, as can be seen in the nonuniform structure that is present in the fit residuals. The FP model fits seem to be more inaccurate that the HW model fits, possibly because the FP region has several well-defined bands, while the HW region has fewer, relatively broad bands. For this reason, accuracy of the model fits in the FP region may be more sensitive to errors in the basis spectra. Basis spectra errors can arise from problems in the spectral data processing, *e.g.* inaccurate background removal, and more importantly, from spectral shifts related to basis spectra sample preparation and environment, as in the β-carotene spectrum.

Fit coefficients for a Raman line image

We can learn more about the accuracy of the basis spectra model by examining the fit coefficients for an entire line image. Fit coefficients for the heterogeneous PIT specimen (UMD016) are shown in Figure 3-5. First we consider the protein fit coefficients. The collagen fit coefficients in both the FP and HW regions exhibit the same trend, decreasing dramatically over the region between $x = 0$ mm to $x = 3$ mm, which closely matches the decrease of blue staining in the histology image. However, the swine fit coefficients are not consistent; the FP coefficient is zero for most of the lesion, while the HW coefficient is approximately 0.2. This discrepancy, combined with the high correlation between the swine and collagen basis spectra in both wavenumber regions, suggests that the swine and collagen fit coefficients are individually inaccurate. Thus, for the remainder of this analysis, we define a single protein coefficient, which is simply the sum of the swine and collagen coefficients.

The triglyceride fit coefficients in both the FP and HW regions remain relatively constant, even in this heterogeneous plaque. This is true throughout the database, suggesting that the triglyceride fit coefficient is not a useful diagnostic measure. Indeed, in biochemical analysis of aortic plaques, Felton *et al.* report that regardless of plaque severity, the triglyceride content remains relatively constant [24]. Thus, while triglycerides are an important contributor to the arterial Raman spectrum, triglyceride content is apparently not a diagnostically relevant parameter, so we do not consider it in the remaining analysis.

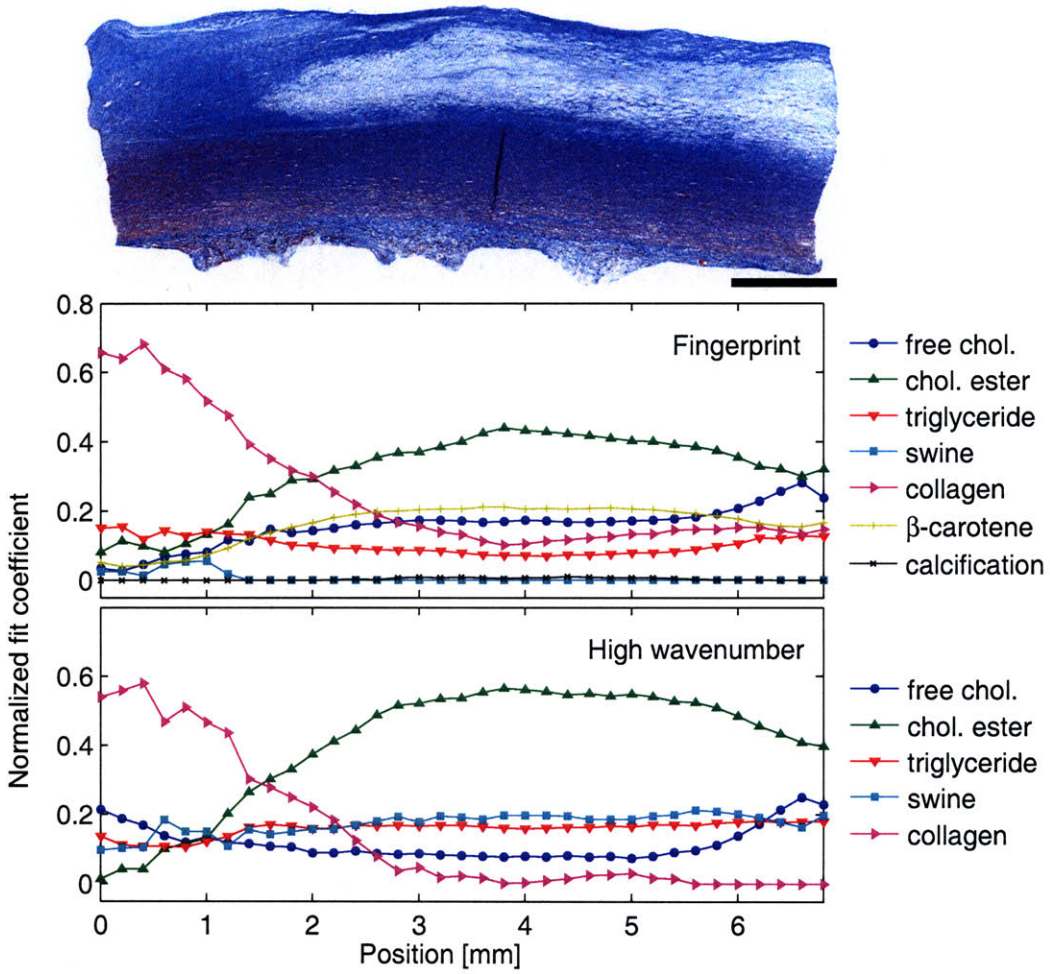


Figure 3-5: FP and HW normalized fit coefficient profiles for a heterogeneous PIT lesion (UMD016). Scale bar, 1 mm.

In the FP region, the β-carotene fit coefficient is highly correlated with the cholesterol ester coefficient, a finding that is also seen throughout the database. This is not surprising as β-carotene binds with cholesterol esters. While β-carotene content is physically correlated with cholesterol ester content, the absolute amount of β-carotene within a plaque strongly depends on the patient's diet [111], and thus β-carotene fit coefficient is not a consistent diagnostic parameter.

By fitting the tissue spectra to our physical basis spectra model, we have reduced the Raman spectral data to only four parameters in the FP region (normalized fit coefficients for free cholesterol, cholesterol ester, protein, and calcification), and three parameters in the HW region (normalized fit coefficients for free cholesterol, cholesterol ester, and protein).

3.2 Plaque characterization based on single site spectra

In this section, we investigate the *ex vivo* Raman database with our physically motivated spectral model, in the context of previously published studies [69, 70, 82, 83]. Mimicking these previous studies, we use spectra taken from a single site to diagnose plaques into three categories: non-lipid plaque, lipid-rich plaque, and calcified plaque.

3.2.1 Histo-pathological classification

A board-certified pathologist assessed the histology slides and identified and diagnosed regions with relatively homogeneous morphology. For each region, the pathologist estimated the percentage of lipid present. A lipid score of 50% was used as a histological threshold for classifying samples as "non-lipid" ($< 50\%$) or "lipid-rich" ($> 50\%$) plaques. To ensure independent sampling, only every third spectral site was used within each region, which separates the spectral sites by 0.6 to 0.75 mm. Forty-six of the 69 specimens were included in this analysis, yielding a total of 121 independent spectral sites, which we term the "single site database." As classified by histology, this single site database included 51 non-lipid sites, 68 lipid-rich sites, and 2 calcified sites.

3.2.2 Fingerprint Raman classification

In the fingerprint region, we use a two-step approach for classification. In the first step, we separate the calcified samples from the non-calcified samples by simply setting a threshold

on the calcification fit coefficient. Other researchers have found that a simple threshold can diagnose calcified vs. non-calcified plaques with sensitivity $> 81\%$ and specificity $> 88\%$ [82, 112]. Because the single site database has only two calcified samples, the threshold was arbitrarily set at 0.4 (Figure 3-6).

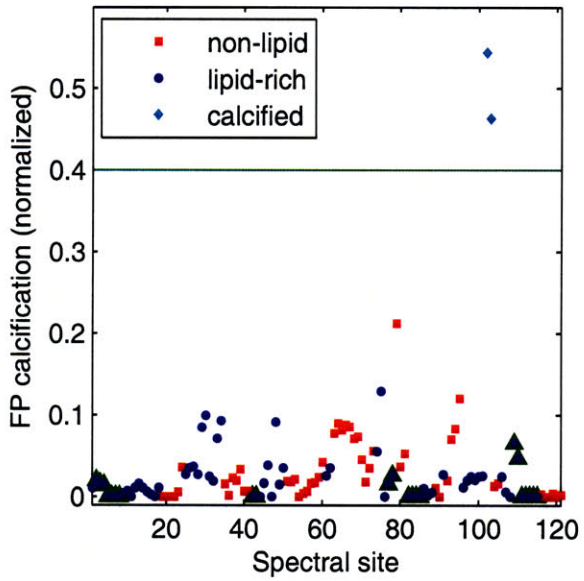


Figure 3-6: In the FP region, a threshold test based on the normalized calcification fit coefficient can be used to classify calcified vs. non-calcified plaques. A subset of the lipid-rich samples were histologically classified as necrotic core; these are highlighted with a green triangle.

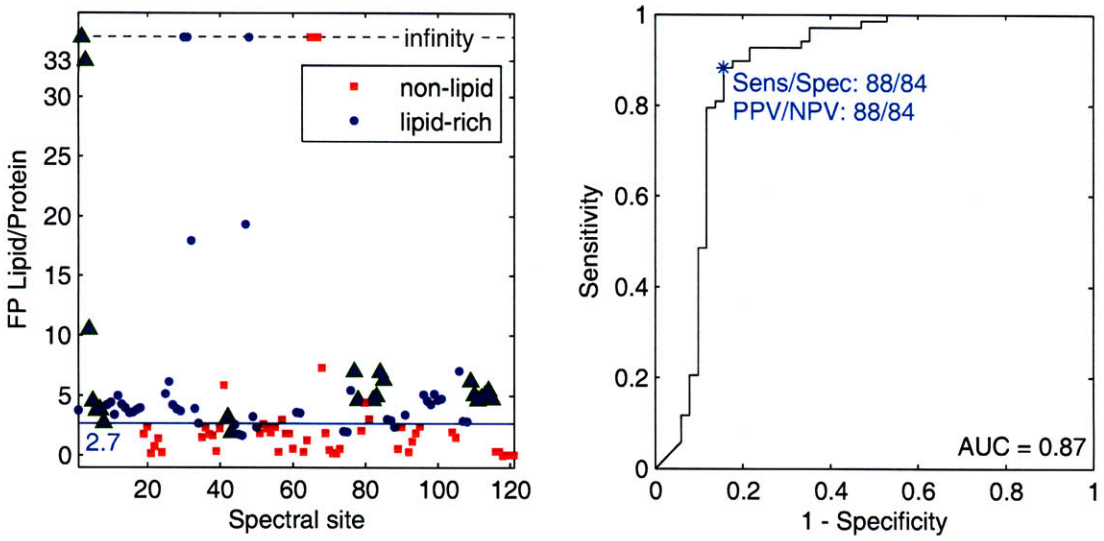
In the second step, the remaining non-calcified plaques were classified into non-lipid and lipid-rich plaques using a threshold test on the ratio of lipid-to-protein. This L/P ratio is defined as

$$\mathrm{L/P} = \frac{\alpha_{\text{free chol.}} + \alpha_{\text{chol. ester}}}{\alpha_{\text{protein}}}. \tag{3.4}$$

Plaques with a L/P ratio greater than the threshold, γ, are diagnosed as lipid-rich. A receiver operating characteristic (ROC) curve was calculated for varying values of γ (Figure 3-7). Choosing $\gamma = 2.7$ results in a sensitivity of 88% and specificity of 84% for identifying lipid-rich plaques. The area under the ROC curve is 0.87.

3.2.3 High wavenumber Raman classification

Because calcified nodules have negligible spectral features in the HW region, we cannot use the HW fit coefficients to diagnose calcified plaques. Based on the histology, the pathologist



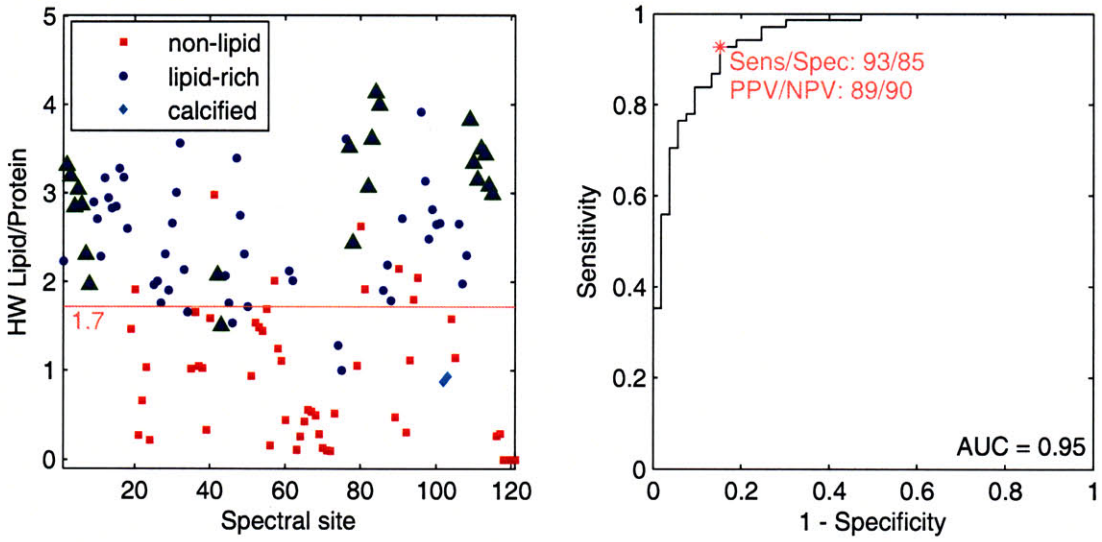
(a) (b)

		Histology		
		Non-lipid	Lipid-rich	Calcified
Raman	Non-lipid	43	8	0
	Lipid-rich	8	60	0
	Calcified	0	0	2

(c) Confusion matrix for $\gamma = 2.7$

Figure 3-7: FP threshold test for classifying non-lipid vs. lipid-rich plaques based on the FP lipid-to-protein ratio. Choosing $\gamma = 2.7$ results in 88% sensitivity and 84% specificity for detecting lipid-rich plaques. A subset of the lipid-rich samples were classified as necrotic core; these are highlighted with a green triangle.

classified the two calcified sites as non-lipid plaques. To classify the samples into non-lipid vs. lipid-rich plaques, we again use the L/P ratio defined in Eq. (3.4), calculated with the HW fit coefficients. An ROC curve was calculated (Figure 3-8), and choosing $\gamma = 1.7$ results in a sensitivity of 93% and specificity of 85% for identifying lipid-rich plaques. The area under the ROC curve is 0.95.



(a) (b)

		Histology		
		Non-lipid	Lipid-rich	Calcified
Raman	Non-lipid	43	5	2
	Lipid-rich	8	63	0
	Calcified	0	0	0

(c) Confusion matrix for $\gamma = 1.7$

Figure 3-8: HW threshold test for classifying non-lipid vs. lipid-rich plaques based on the HW lipid-to-protein ratio. Choosing $\gamma = 1.7$ results in 93% sensitivity and 85% specificity for detecting lipid-rich plaques. A subset of the lipid-rich samples were classified as necrotic core; these are highlighted with a green triangle. The two calcified sites are marked in the plot, but for this HW analysis, they were included in the non-lipid group.

3.2.4 Discussion

Both the fingerprint and high wavenumber regions are capable of identifying non-lipid vs. lipid-rich plaques, with sensitivity and specificity similar to previous studies (Table 1.1). The HW diagnostic algorithm performs slightly better than the FP algorithm, suggesting that the HW spectral model is more accurate for distinguishing lipid vs. protein bases.

This may be due to the higher degree of orthogonality between the lipid and protein bases in the HW region vs. the FP region.

In the FP region there are seven sites with $(L/P)_{FP} = \infty$, because the protein fit coefficient for these sites is zero. Closer inspection of the spectra at these seven sites reveals that one suffers from low SNR because of high autofluorescence from the tissue, three have low SNR because the lesion is thin and protein-rich, resulting in low Raman signal intensity, and three sites exhibit an unusual background signal that was not accurately subtracted by the background removal algorithm. These sites could be automatically excluded from the database analysis by using a metric based on SNR; these sites could then be termed "non-diagnostic." The HW region does not exhibit similar problems, because HW spectra have higher SNR and because background removal is easier in the HW region. Indeed all seven of these samples are classified correctly in the HW diagnostic algorithm.

One disadvantage to the HW region is that it cannot detect the presence of calcifications in the atherosclerotic lesion. By design, our *ex vivo* Raman database has a small number of calcified lesions, so this does not affect the overall accuracy of the HW algorithm. However, the HW region's lack of sensitivity for calcification may be problematic for studying atherosclerosis *in vivo*.

It may be possible to develop a classification algorithm that uses the fit coefficients in both the FP and HW regions simultaneously, potentially improving the classification accuracy. For example, a combined algorithm might first leverage the FP region's sensitivity to calcification to classify the calcified plaques, and then use a weighted combination of the FP and HW L/P ratios, with a higher weighting on the HW ratio to reflect its robustness to low SNR.

3.3 Raman-histology spatial correlations

Compared to single site classification, analyzing Raman spectra over an extended spatial region provides a more complete characterization of hetergeneous atheroclerotic lesions. In this section, we consider the normalized fit coefficients for complete Raman line images.

A heterogeneous PIT specimen (UMD016) is shown in Figure 3-9. Based on histological evaluation, the plaque is primarily protein-rich between $x = 0$ mm and $x = 1.4$m, where the beginning of a lipid-rich region can be seen. Between approximately 1.4 to 3.4 mm, the

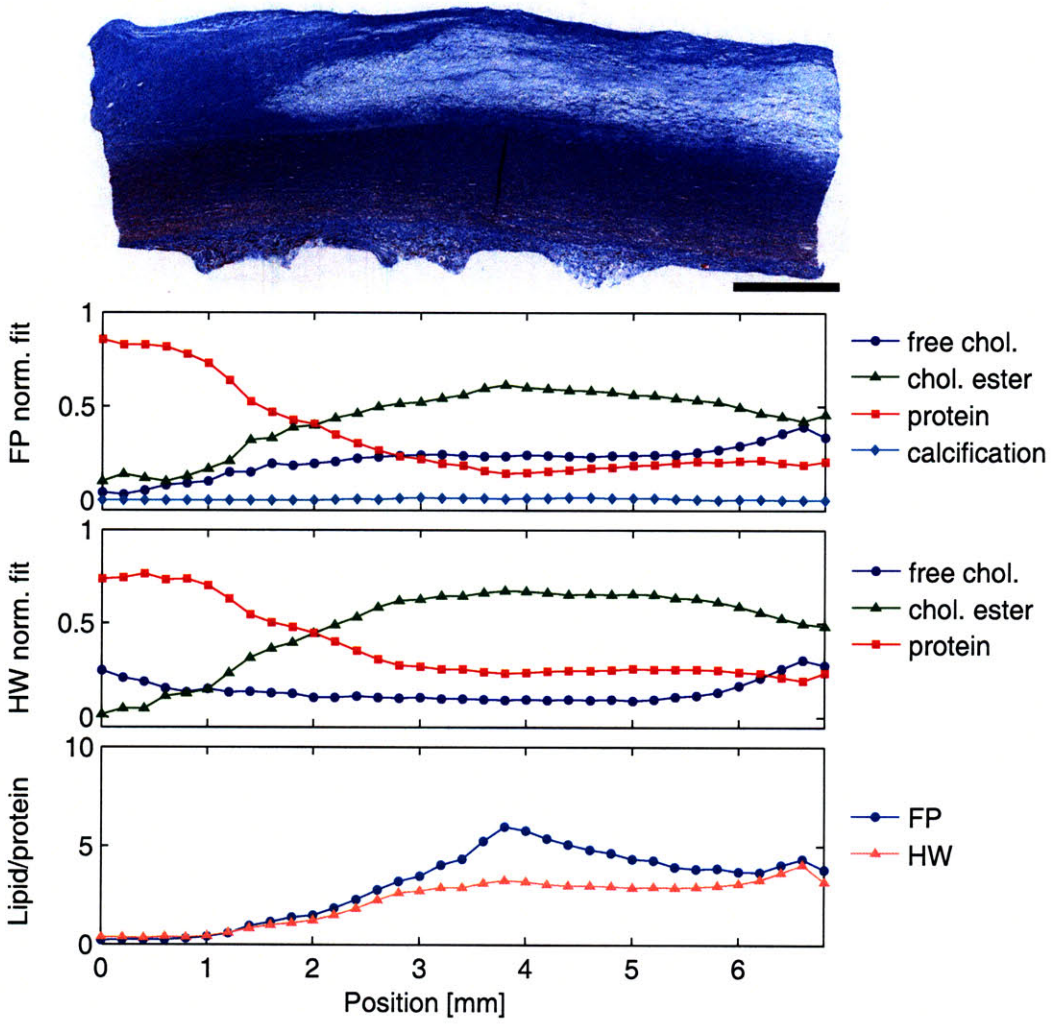


Figure 3-9: A heterogeneous PIT lesion (UMD016), with normalized FP and HW fit coefficient profiles. FP and HW lipid-to-protein ratios are shown in the bottom panel. Scale bar, 1 mm.

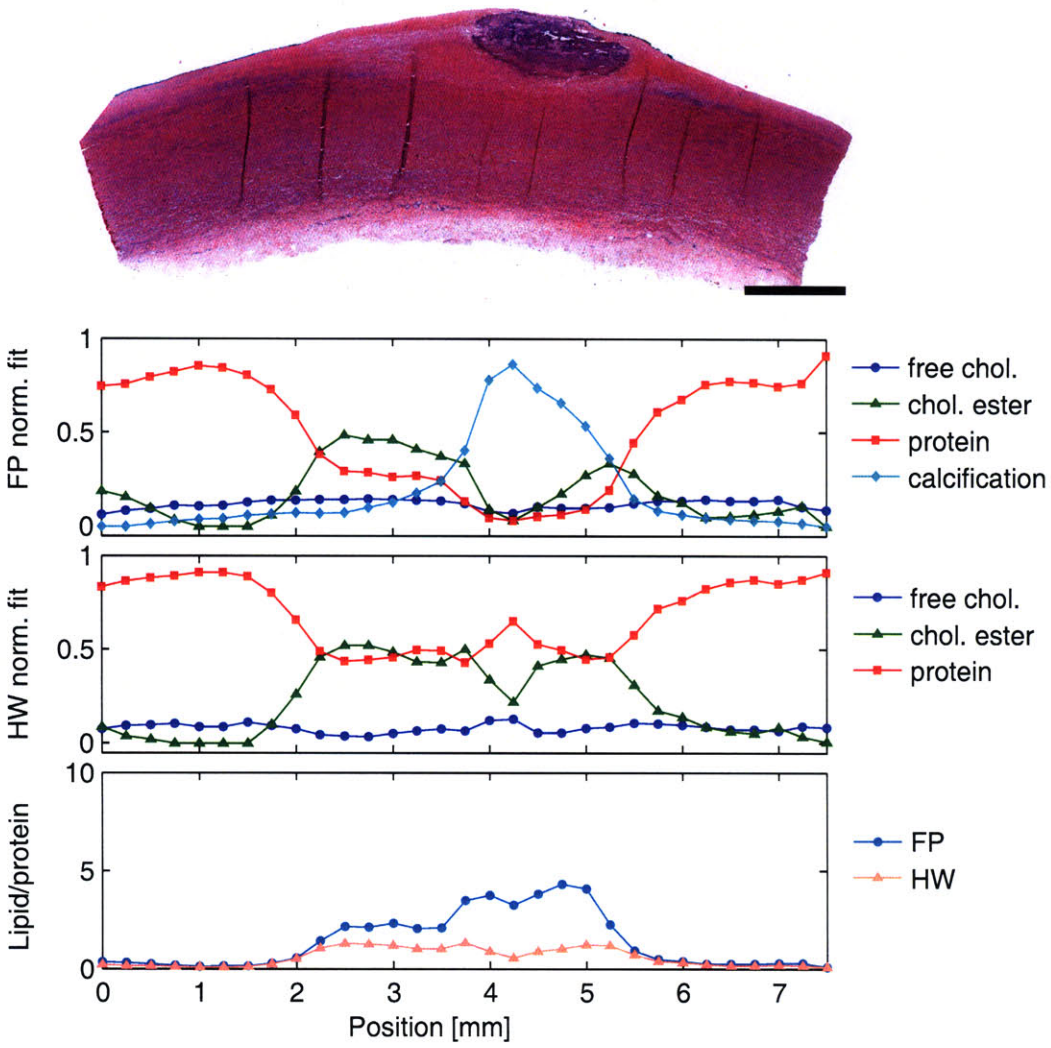


Figure 3-10: A calcified plaque, with normalized FP and HW fit coefficient profiles. Lipid-to-protein ratios are shown in the bottom panel. Scale bar, 1 mm.

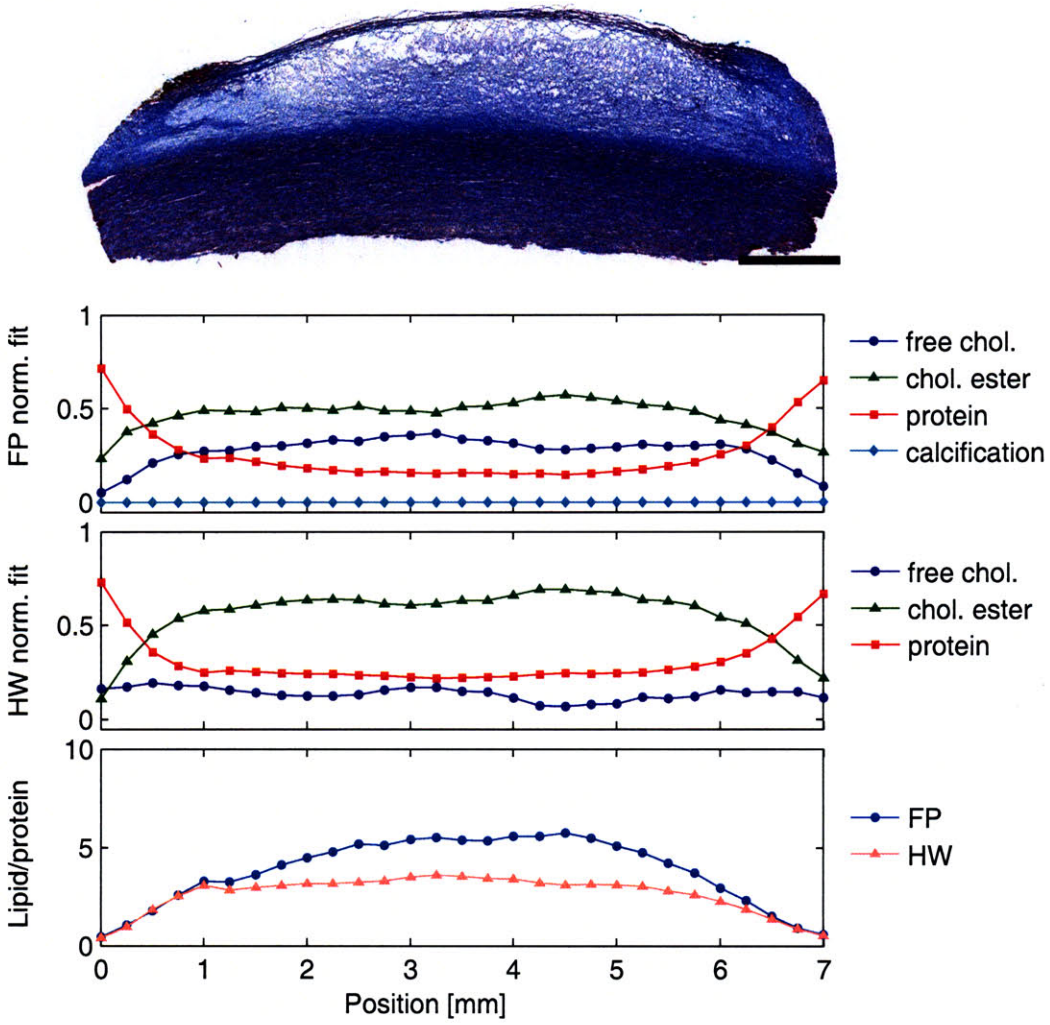


Figure 3-11: A necrotic core plaque, with normalized FP and HW fit coefficient profiles. Lipid-to-protein ratios are shown in the bottom panel. Scale bar, 1 mm.

amount of lipid present increases, while the protein content gradually decreases. For the remainder of the section, from 3.4 mm to the end (6.8 mm), the specimen is primarily lipid-rich, but not necrotic. Spectrally, the same features are seen in the normalized fit coefficient profiles. In the FP region, the protein coefficient starts at 0.86 at $x = 0$ mm and remains relatively constant for the first 0.6 mm, before gradually decreasing to 0.2 at $x = 3.2$ mm and remaining constant for the remainder of the sample. Both the free and esterified cholesterol fit coefficients increase with increasing position. However, free cholesterol reaches a plateau around $x = 3$ mm while the cholesterol ester fit contribution continues to increase. The FP calcification fit coefficient is zero throughout the sample. Fit coefficients in the HW region follow the same general trends as in the FP region. There is a discrepancy in the relative cholesterol coefficients between 0 and 1 mm, but this occurs where the protein fit coefficient dominates the spectrum, which makes the remaining fit coefficients unimportant. Both the FP and HW L/P ratios correspond well with the histology.

An example of a calcified specimen (UMD045) is shown in Figure 3-10. As expected, the calcification fit coefficient is very high in the purple, calcified region ($x = 3.75$ to $x = 5.25$ mm). At the edge of the calcified nodule, a modest amount of lipid is seen in the histology section between 2.5 mm and 3.75 mm, which corresponds well with an increase in the cholesterol ester fit coefficient seen in both the FP and HW regions. The cholesterol fit is relatively constant throughout the sample. Aside from regions where the protein fit coefficient is very high, the free cholesterol and esterified cholesterol fit coefficients correspond very well between the FP and HW regions. The protein coefficient in the HW region spikes at 4.25 mm, in the calcified region, due to the lack of calcification signal in the HW region. Both the FP and HW L/P ratios are elevated over the region of the plaque where the intima is thickened.

In the necrotic core specimen (UMD087) of Figure 3-11, the protein fit coefficient accurately represents the increased protein composition at the shoulders of plaque. As expected, the L/P ratios are elevated over the lipid pool area. The behavior of the free cholesterol fit coefficients is not consistent between the FP and HW regions.

Another heterogeneous PIT specimen (UMD089) is shown in Figure 3-12. The FP and HW L/P ratios are consistent with each other, and they suggest that there is elevated lipid between $x = 1$ and $x = 3$ mm. However, only a very mild increase in lipid content is seen in the corresponding region of the trichrome histology section. Examination of the gross

pathology photograph (Figure 3-12(b)) shows a yellow, lipid-rich region just offset from the histology section plane. The Raman sampling volume of the benchtop system used in this study is several hundred microns in diameter, while the histology section is only 5 μm thick, so it is likely that the elevated L/P ratio results from a lipid region of this heterogeneous plaque that lies outside of the region shown in the histology section.

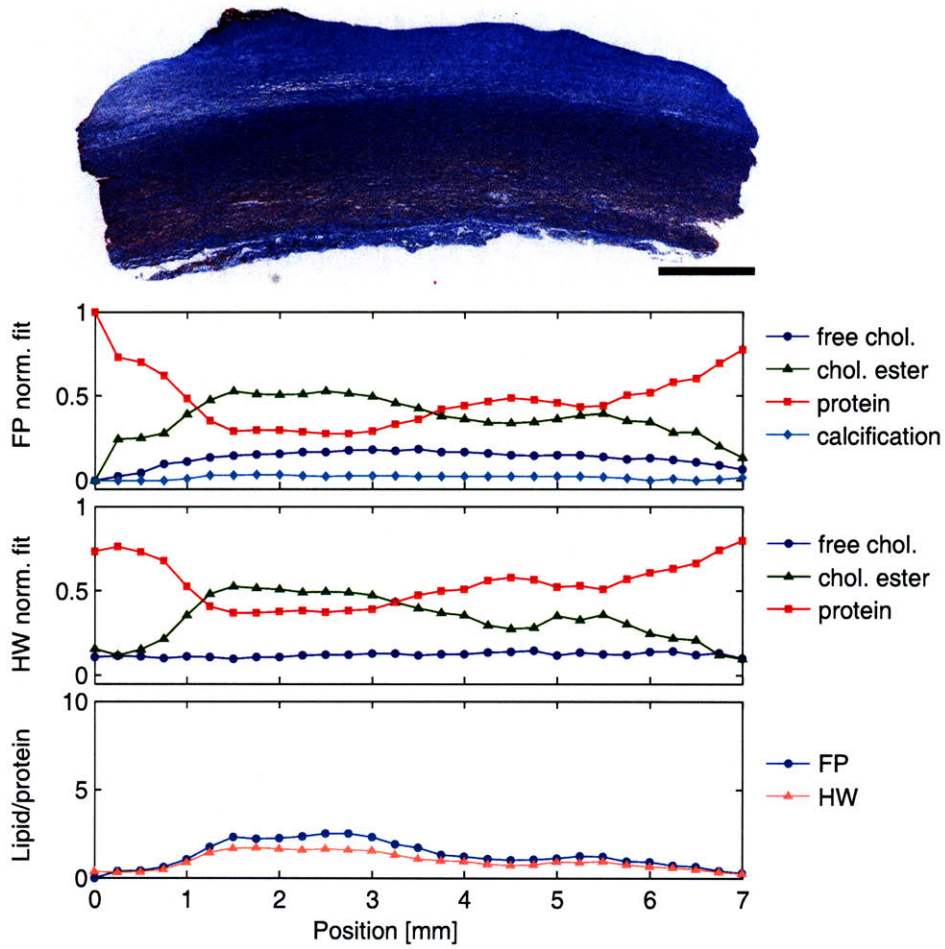
3.3.1 Discussion

In general, the FP and HW normalized fit coefficients consistently exhibit the same trends, correlating well with the overall lipid and protein content of the lesions, as evaluated through histology. In all specimens, the L/P ratios are also consistently highly correlated between the FP and HW regions, confirming that both the FP and HW basis spectra models are capable of determining the relative fit contributions of cholesterols and proteins. The FP L/P ratio is generally greater than or equal to the HW L/P ratio, a finding that is seen in all the examples shown, as well as throughout the remaining specimens in the database. This apparent scaling between the two ratios may be due to differences in the Raman cross sections between bands in the FP and HW regions, or possibly to inaccuracy in the spectral model or in the fitting procedures.

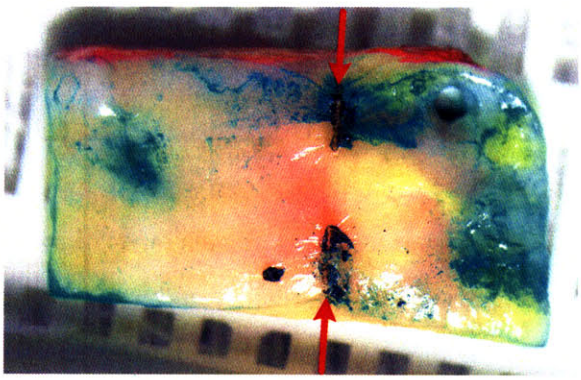
For free cholesterol, the FP and HW fit coefficients are not consistent, with the HW model most often underestimating the contribution of free cholesterol relative to the FP model. While this may be explained by differences in the Raman cross section, the high correlation between the basis spectra for free cholesterol and cholesterol ester in the HW region ($\rho = 0.97$) suggests that the discrepancy is most likely explained by difficulties in the model fitting procedures. Currently, for the specimens in the database, we have no external gold standard for determining free and esterified cholesterol content, so we cannot assess the accuracy of either region for making this determination.

3.4 Discussion

In this chapter, we have developed a physically motivated set of basis spectra to model the tissue Raman spectra. The model includes the most important chemicals that contribute to the Raman spectrum of arterial tissue: free cholesterol, cholesterol ester, triglyceride, and proteins, and in the FP region only, β-carotene and calcification. The lipid



(a)



(b) Photograph

Figure 3-12: Normalized fit coefficient profiles (a) for a heterogeneous lesion diagnosed as pathological intimal thickening (UMD089), with corresponding gross pathology photograph (b). The lipid-to-protein ratio suggests elevated lipid between 1 to 3 mm, which is barely discernible in the histology image. Red arrows in the photograph denote the ends of the Raman scan lines and the presumed location of the histology section. Yellow, lipid-rich material is present just to the right of the red arrows, suggesting that the histology section barely missed the lipid-rich region. Scale bar, 1 mm.

and β-carotene spectra were acquired from commercially purified chemicals: cholesterol, cholesteryl linoleate, and glyceryl trioleate. Protein spectra were acquired from two sources: *ex vivo* swine aorta, which is predominantly elastin and collagen, and tissue-engineered cartilage, which is primarily collagen. The calcification spectrum was acquired by dissecting a large calcified nodule from an *ex vivo* human aorta. Actin, phospholipids, and other cholesterol esters were not separately represented in the model, which may affect the accuracy of the model.

The tissue Raman spectra are modeled as a linear combination of the basis spectra, and the weighting coefficients for each basis are calculated using a constrained least squares algorithm. The basis spectra are linearly independent, so the weighting coefficients can be uniquely determined. However, confidence in the physical interpretation of the weighting coefficients is related to the orthogonality of the bases. For example, consider two linearly independent vectors, $\mathbf{v_1}$ and $\mathbf{v_2}$. The vector $\mathbf{z} = \mathbf{v_1} + \mathbf{n}$, where $\mathbf{n}$ is a noise vector in the span of $\mathbf{v_1}$ and $\mathbf{v_2}$, can be decomposed into $\mathbf{v_1}$ and $\mathbf{v_2}$ uniquely. However, if $\mathbf{v_1}$ and $\mathbf{v_2}$ are highly correlated, only a small noise vector, $\mathbf{n}$, is needed to cause $\mathbf{z}$ to look more like $\mathbf{v_2}$ than $\mathbf{v_1}$. Thus the weighting coefficients for highly correlated basis spectra are not robust to perturbations in the tissue spectrum, and the physical interpretation of these coefficients, relative to each other, may be suspect.

For lesion characterization, we are concerned with distinguishing the protein group from the lipid group and with distinguishing free cholesterol and cholesterol ester. Between proteins and lipids, the spectral model in the HW region is more orthogonal ($\rho = 0.53$) than the FP region ($\rho = 0.68$), which suggests that the HW region is better for distinguishing lipid from protein, a finding that is also confirmed by the database data. For cholesterol and cholesterol ester, the HW basis spectra are far more highly correlated ($\rho = 0.97$), than the FP region ($\rho = 0.81$), suggesting that the HW spectral model will have more difficulty distinguishing these two components than the FP model.

Both the FP and HW basis spectra models are capable of reconstructing the tissue spectra in the *ex vivo* Raman database. The models accurately account for all the significant bands in the tissue spectra, though in some cases the peak intensities are not accurate, possibly due to problems with the data processing procedures and the incompleteness of the spectral models. Nevertheless, the fit coefficients do accurately reflect observations in the histology sections; coefficients for proteins, lipids, and calcification correlate well with

the morphology seen in the histology sections.

A subset of the spectral model fit coefficients were used for lesion characterization. Free cholesterol and cholesterol ester were retained because of their importance in determining disease state. The swine and collagen fit coefficients were summed into a single protein coefficient. In the FP region, the calcification fit coefficient was also retained. Triglycerides and β-carotene are not considered to be diagnostically relevant and thus were not used for lesion characterization.

The database data was analyzed in two ways. First, we developed a diagnostic algorithm based on single site analysis, which is analogous to previously published studies. Independent sites were chosen from the database specimens and histologically classified into non-lipid, lipid-rich, and calcified groups. We then used the Raman fit coefficients to diagnose the sites into those categories. In the FP region, this was accomplished in two stages. In the first stage, the calcified samples were classified by using a threshold test on the calcification fit coefficient. The remaining non-calcified samples were then classified by thresholding the L/P (lipid-to-protein) ratio. Since the HW region has no calcification signal, only the L/P ratio was used. For detecting lipid-rich plaques, the L/P threshold test has sensitivity and specificity of 88% and 84% in the FP region and 93% and 85% in the HW region. These results are similar to previously published studies.

Second, the Raman line images were analyzed. Analyzing the fit coefficients over the entire line image allows for better characterization than examining single site spectra. Even without further data processing or a classification algorithm, it is easy to visually identify the presence of a lesion and estimate its size based on the normalized fit coefficients alone.

Both the FP and HW regions are capable of detecting lipid-rich plaques based on the L/P ratio, but the HW region performs slightly better for several reasons. First, HW Raman signals have higher SNR, due to the Raman cross sections in the HW region being larger than in the FP region and due to the presence of less fluorescence. Second, the HW spectra are easier to fit to the spectral model. The HW spectra have few, relatively broad bands, so the spectral model fitting procedure is more robust to errors in both background removal and in determining the basis spectra. Background removal is also easier in the HW region than in the FP region, due to decreased fluorescence and the HW region's relatively narrow bandwidth. Finally, the basis spectra for lipids and proteins are more orthogonal in the HW region, increasing the accuracy of the L/P ratio.

In contrast, the FP region is more susceptible to spectral model fitting errors. As noted in Section 2.2.1, background removal is more challenging in the FP region. The FP region is also more likely to have basis spectra errors, which may arise from spectral shifts related to basis spectra sample preparation or from poor background removal in the basis spectra. However, the FP region has a significant advantage over the HW region in being able to identify calcified regions of the plaque. The FP calcification basis spectrum is orthogonal to all the other basis spectra. Combined with the large Raman cross section for PO_4^{-3}, this makes calcification easy to distinguish, even in minute amounts.

An ideal extension of Raman spectroscopy-based lesion characterization lies in determining the relative composition of the cholesterols, *i.e.* distinguishing between relative contributions of free cholesterol vs. cholesterol esters, and if possible, between individual cholesterol esters, particularly cholesteryl linoleate and cholesteryl oleate. As seen in Figure 3-1, the spectral differences between the cholesterol and cholesteryl linoleate basis spectra are subtle, but the FP region, with $\rho = 0.81$ for these bases, is more likely to be able to make this distinction than the HW region ($\rho = 0.97$). Without having validated the data processing and spectral modeling procedures, we cannot confidently claim that our current spectral model is capable of making this distinction. Using Raman spectroscopy to analyze cholesterol content will require further development, particularly in creating basis spectra that accurately represent Raman activity of the cholesterol constituents *in situ* and in carefully validating the preprocessing, background subtraction, and spectral decomposition procedures.

External validation of the Raman spectroscopy results is essential to further developing this technique. In particular, the data processing steps, including background subtraction, and the spectral modeling need to be validated with a method other than histology. While Raman spectroscopy is sensitive to molecular composition and molecular structure, histology assesses the microscopic organization and orientation of the cellular structure. It is not clear that there is a one-to-one correspondence between a lesion's molecular composition and cellular structure, especially at different time points during lesion development.

An alternate validation method, which is left for future work, is to correlate the Raman spectra and fit coefficients with chemical composition via biochemical assays, similar to studies published by Baraga and Brennan *et al.* [67, 68]. However, biochemical assays require chemical processing of a relatively large volume of tissue (> 1 mm^3), which may

preclude direct validation of the Raman line images.

An important limitation of the current study is the use of two-dimensional histology sections to correlate with the three-dimensional Raman sampling volumes. Raman spectroscopy provides a bulk measurement with a sampling volume on the order of 1 mm^3. Due to the heterogeneous nature of atherosclerotic lesions, a single thin histology section cannot completely characterize the morphology of a three-dimensional Raman sampling volume. This limitation can be overcome in future studies by obtaining serial histology sections throughout the Raman sampling volume.

The use of histology as the gold standard for lesion diagnosis imposes another limitation on the development of Raman spectroscopy for diagnosis. Histological sections are used by pathologists to diagnose lesions into discrete categories, but a limited number of discrete categories may not be able to adequately characterize the wide range of atherosclerotic lesion types typical of a complex disease. Furthermore, histological sections provide a snapshot of lesion morphology at a single point in time, but it is possible that lesions at different stages of development may look morphologically similar while having different chemical signatures. Raman spectroscopy, being sensitive to both molecular composition and molecular structure, may be able to detect pathological changes in tissue before they are visible in histology.

3.5 Conclusions

In this chapter, we have developed a physically motivated basis spectra model to decompose the Raman spectra into lipid, protein, and calcification constituents. After fitting the data to the spectral model, we classified single site Raman spectra into non-lipid and lipid-rich categories, by thresholding the ratio of the lipid fit coefficient to the protein fit coefficient. In both the fingerprint and high wavenumber regions, the resulting sensitivity and specificity for detecting lipid-rich plaque is comparable to previously published fingerprint Raman studies. By looking at fit coefficient profiles for Raman line images, it is possible to determine the presence and size of an atherosclerotic lesion. In terms of diagnostic capability, only the fingerprint region is able to identify calcifications, but both the FP and HW regions can distinguish between non-lipid and lipid-rich plaques.

Chapter 4

Raman Probe

Intracoronary Raman spectroscopy requires the use of a flexible, small diameter (~1 mm) fiber optic probe. The need for two separate channels with distinct optical filters at the catheter's distal tip presents a significant challenge for Raman probe fabrication. While several FP Raman probes have been demonstrated for atherosclerosis, most have not been suitable for intracoronary Raman spectroscopy because they were either too large and inflexible to be introduced into human coronary arteries or have not had sufficient optical throughput. In this chapter, we describe a prototype Raman probe, a miniaturized version of the previously demonstrated Raman probes that is capable of acquiring spectra in both the fingerprint and high wavenumber regions.

4.1 Design and fabrication

4.1.1 Fiber bundle

One way to increase the signal collected by the probe is to have as large a collection area as possible, through the use of large diameter multimode fibers. However, larger diameter fibers are stiffer than smaller diameter fibers, decreasing probe flexibility. For this reason, our probe uses 100 μm core diameter fibers, which provide a tradeoff between large area and small bend radius. Similar to the previously demonstrated probes, our design consists of a ring of six collection fibers arranged around a single excitation fiber. All fibers were identical, low-OH silica-silica fibers, with an acrylate buffer (Polymicro Technologies; Phoenix, AZ). The fibers had a 120 μm cladding diameter, which enabled the use of standard off-the-

shelf fiber optic components and tools. The acrylate buffer diameter was 160 μm and the numerical aperture of the fiber was 0.26.

4.1.2 Optical filter

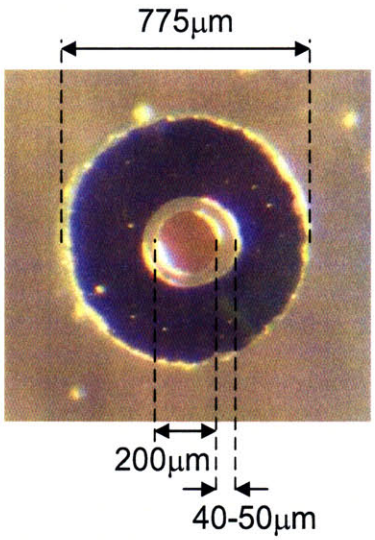
As described in Section 1.2.1, one of the challenges in creating a fingerprint Raman probe is the need for optical filtering at the distal tip of the probe to reduce the amount of background signal generated in the fiber. The traditional approach is to filter the excitation fiber with a shortpass or bandpass filter, which blocks the fiber background from entering the tissue. A longpass filter is placed over the collection fibers, which allows only the Raman and fluorescence signals from the tissue to propagate back to the spectrometer. Blocking the elastically scattered excitation light from entering the collection fibers prevents the generation of an additional fiber background signal. Dielectric interference filters, composed of multiple thin films of dielectric material deposited onto a fused silica substrate, can achieve high transmission and high extinction with a sharp transition edge, and are an ideal solution for Raman spectroscopy. However, the need for two separate optical filters in a small diameter (~1 mm) probe has been a significant engineering challenge for intracoronary Raman probe development.

Previous researchers have fabricated Raman probes utilizing filters deposited on two or more separate substrates. Shim *et al.* cut individual filters for each of the eight fibers in their probe, with each filter being a ~300 μm diameter disk [94]. Motz *et al.* simplified the fabrication process by using a combination of two separate filters, a central disk and an outer annulus. The shortpass filter disk had a 500 μm diameter and was 1 mm thick. The annular longpass filter had an inner diameter of 700 μm and outer diameter of 1.75 mm [86]. Both of these probes are unsuitable for intracoronary use because they are not flexible enough and are too large in diameter.

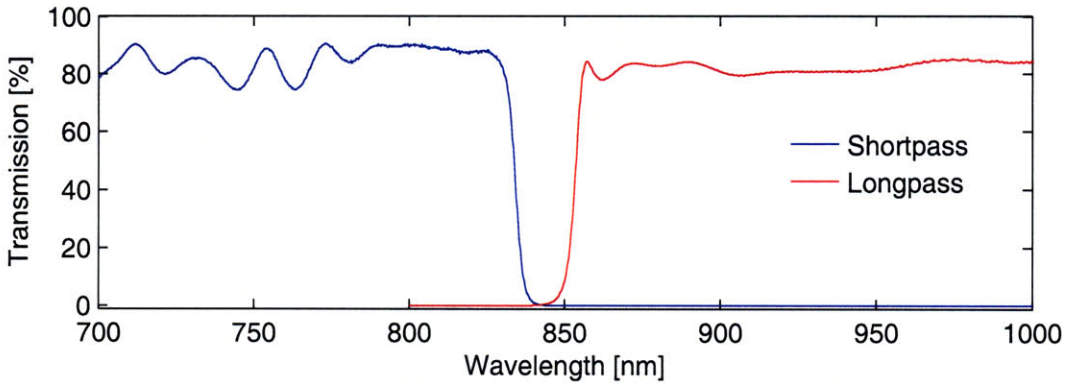
Using a similar approach of a central shortpass filter disk and an annular longpass filter, Komachi *et al.* demonstrated a Raman probe with much smaller dimensions. The central filter was 150 μm in diameter, with a thickness of 150 μm. The annulus had inner and outer diameters of 200 μm and 500 μm, and the thickness was 100 μm [95]. The fabricated Raman probe was only 600 μm in diameter, but fabrication of the filters at this size scale was a technically challenging process, requiring specialized equipment and expertise.

We have circumvented this difficulty by using photolithography to deposit the shortpass

and longpass filters directly onto a single fused silica substrate (BARR Associates; Westford, MA). The central region of the filter is a 200 μm diameter shortpass filter with a cutoff wavelength of 830 nm (Figure 4-1), while the outer annular region is a longpass filter with a cutoff of 860 nm. Because of current limitations in the photolithography process, there is a $40 - 50$ μm dead space between the two filter regions. The outer diameter of the filter is 775 μm, and the thickness of the fused silica substrate is 600 μm.



(a) Photograph of the patterned filter surface



(b) Filter transmission

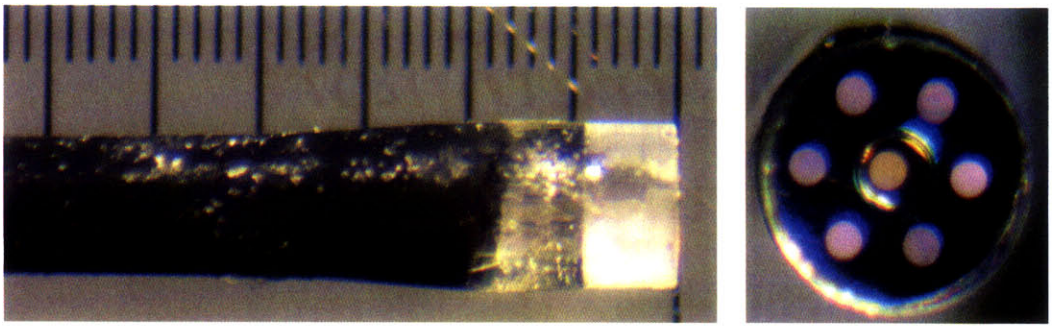
Figure 4-1: Patterned filter. In (a) the longpass filter is deposited in the outer annulus while the shortpass filter is located in the center. The gap between the two filters is clearly visible.

The shortpass and longpass filters have excellent spectral response (Figure 4-1). The shortpass filter has transmission of 83% at the FP excitation wavelength (830 nm), and 79% at the HW excitation wavelength (740 nm). The shortpass filter effectively blocks the

fiber background with transmission less than 0.05% in the wavelength range 850 – 1000 nm. The outer annulus longpass filter transmits the returning Raman light (between 866 – 1100 nm) with transmission greater than 80%, and transmits less than 0.01% of the elastically scattered light to the collection fibers.

4.1.3 Probe assembly

To assemble the fiber bundle, the seven fibers were first cleaved to provide a flat fiber face. The fiber faces were aligned and bundled together using medical grade heat shrink tubing (Advanced Polymers Incorporated; Salem, NH). Because the fibers are identical in diameter, they self-assemble into a 6-around-1 configuration. However, to avoid conflict with the dead space of the filter, the outer ring of fibers was deliberately spaced farther from the central fiber than necessary. To retain the self-assembling property, a short length (~1 cm) of heat shrink tubing was placed at the distal end of each fiber to increase the effective fiber diameter in the bundled region. In the final assembled bundle, the center-to-center distance between the fibers is approximately 200 μm.



(a) Side view **(b)** Front view

Figure 4-2: Prototype Raman probe. The forward-viewing Raman probe consists of a fiber bundle bonded to a patterned dielectric filter. In (a), the filters are bundled together with black heat shrink tubing and epoxied to the filter. In (b), the fibers have been illuminated with overhead room lights. The central fiber appears orange due to the shortpass filter, while the collection fibers appear purple due to the longpass filter. The outer diameter of the filtered bundle is 775 μm. Rule marks in (a) are 100 μm apart.

The fiber bundle was then epoxied to the filter, with the thin film surface in contact with the fiber face. To align the fiber bundle with the filter, both the filter and the fiber bundle were mounted in a nanopositioning stage (Thorlabs; Newton, NJ), and a helium-neon laser was coupled to the proximal end of the excitation fiber. The He-Ne beam is reflected by the annular longpass filter, so by visualizing the assembly under a microscope, it was possible

to center the excitation fiber over the shortpass filter. Medical grade epoxy (Loctite Hysol M-31CL, Henkel; Dusseldorf, Germany) was used to bond the fiber bundle and filter (Figure 4-2).

Proximal to the heat shrink tubing, the fibers were approximately two meters long. The Raman probe was modeled after commercially available intracoronary catheters, which have a working length of ~1.5 m. Thus, 1.5 m of the distal portion of the the optical fibers were loosely bundled together in plastic tubing. The proximal 0.5 m of fiber was split into two channels; the single excitation fiber was coupled to the lasers and the six collection fibers were coupled to the spectrometer. Both sets of fibers were inserted in fiber optic furcation tubing for protection.

The outer diameter of the forward-viewing Raman probe is 775 μm and is currently limited by the outer diameter of the filter module. The fiber bundle diameter is ~560 μm, though this could be reduced by placing the outer ring of collection fibers closer to the central excitation fiber and re-patterning the filter accordingly. By re-engineering the filter and fiber bundle, the Raman probe could be made as small as ~500 μm.

4.1.4 Coupling to the Raman system

The excitation fiber was connectorized with an FC connector and coupled to both the 830 nm and 740 nm laser, in a manner similar to the method described in Section 2.1.1. To couple the collection fibers into the spectrometer, the proximal tips of the fibers were arranged in a linear array, epoxied together, polished, and then aligned at the input of the spectrometer. This geometry maps each fiber to a different vertical position on the CCD (Figure 4-3). The 100 μm core diameter matches the internal slit of the benchtop Raman system, yielding a spectral resolution of 8 cm^{-1}.

4.1.5 Benchmarks

Probe background

An important metric for evaluating the quality of a Raman probe is based on the magnitude of the self-generated probe background signal. While the most important component of the probe background is typically the Raman signal generated within the silica fibers, the background signal may also arise from other materials within the optical path, such as

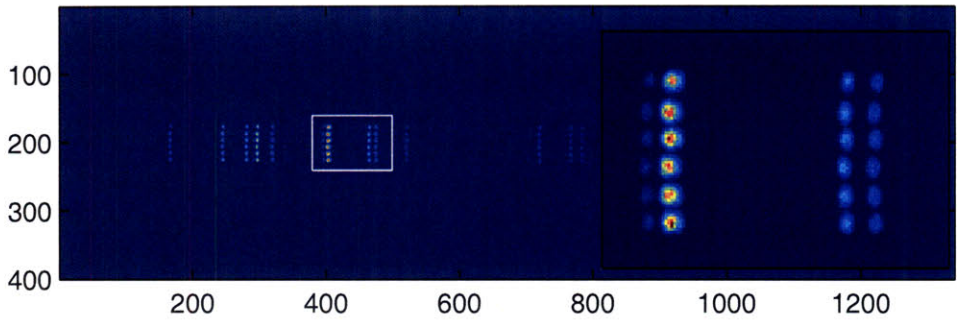


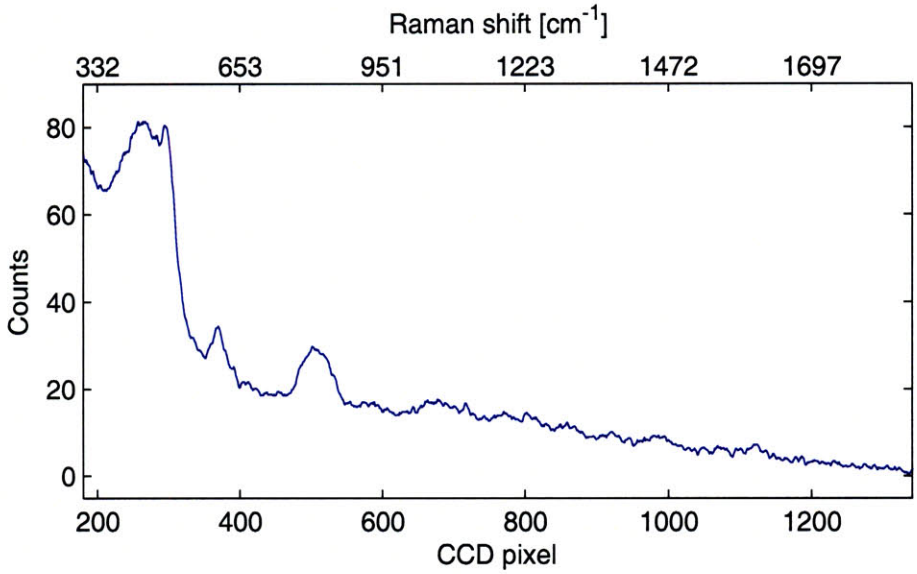
Figure 4-3: The Raman probe was coupled to the spectrometer and illuminated with a neon lamp. The individual fibers can be seen in the resulting CCD intensity image. Numbers on the axes are CCD pixel numbers; incoming light is dispersed in wavelength along the x-axis. The inset shows an expanded view of the fiber alignment.

epoxy and tubing. The shape and magnitude of the probe background can be assessed by acquiring the Raman spectrum while the probe is held in air, not interrogating any sample. The background signal for a representative Raman probe is shown in Figure 4-4. First, the background spectra were acquired with the FP or HW laser on and the catheter held in air. Then the lasers were blocked and a dark spectrum was acquired in ambient light conditions. The ambient spectrum was subtracted from each background spectrum, and the resulting spectrum was smoothed with a Savitzky-Golay filter. This is a best case scenario; when the probe is illuminating a scattering sample, a higher magnitude background signal may result.

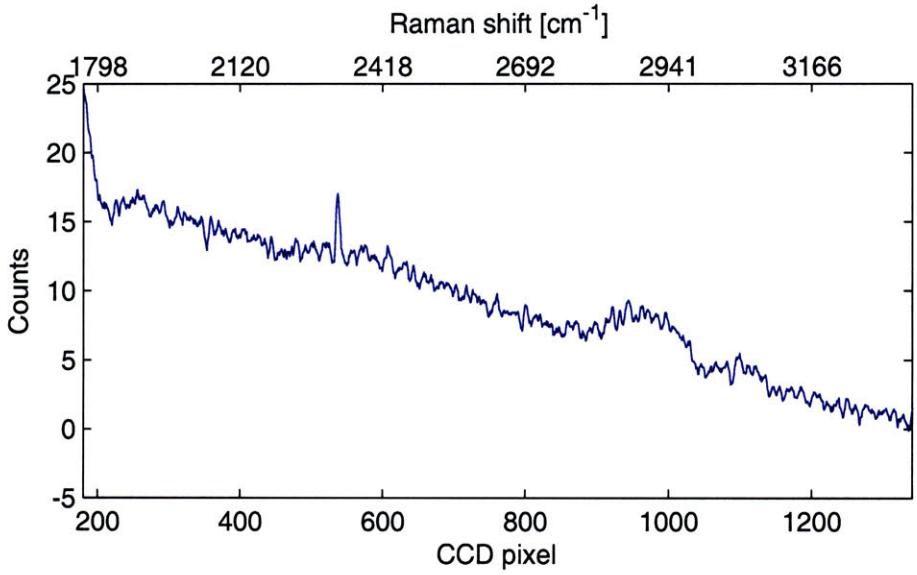
The FP background spectrum exhibits the classic silica fiber Raman background [113], with characteristic bands at $300 - 530$ cm^{-1}, at 603 cm^{-1}, and at 809 cm^{-1}. The FP background signal intensity is very low, with a maximum of only 80 counts. The HW background spectrum has few features. In the HW region of interest ($2818 - 3021$ cm^{-1}), there is a small, broad background signal, with a magnitude of only $\sim$10 counts. The small magnitude of the background signals in both wavenumber regions demonstrates both the effectiveness of the filter and the good fabrication of this particular probe.

Acetaminophen spectrum

Raman spectra of acetaminophen powder acquired with the probe exhibit all of the expected peaks without distortion (Figure 4-5). The FP spectrum has a slight silica Raman signal

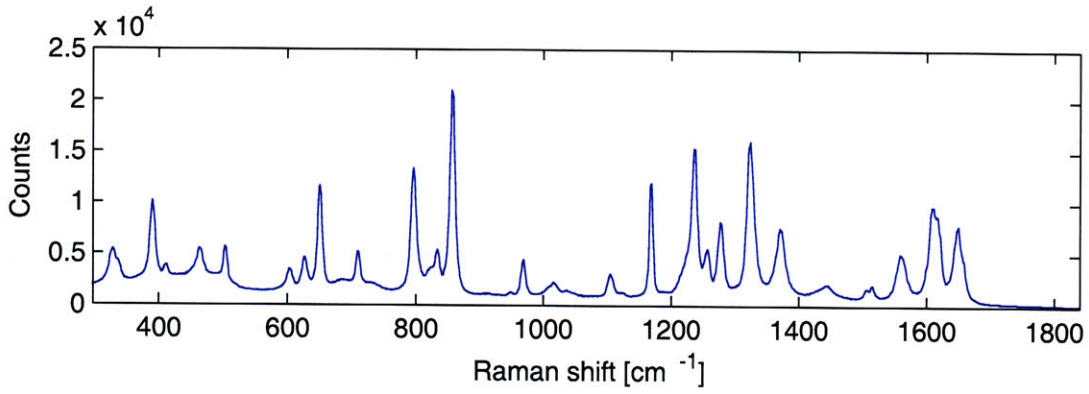


(a) Fingerprint background spectrum (100 mW, 60 frames @ 0.25 s/frame)

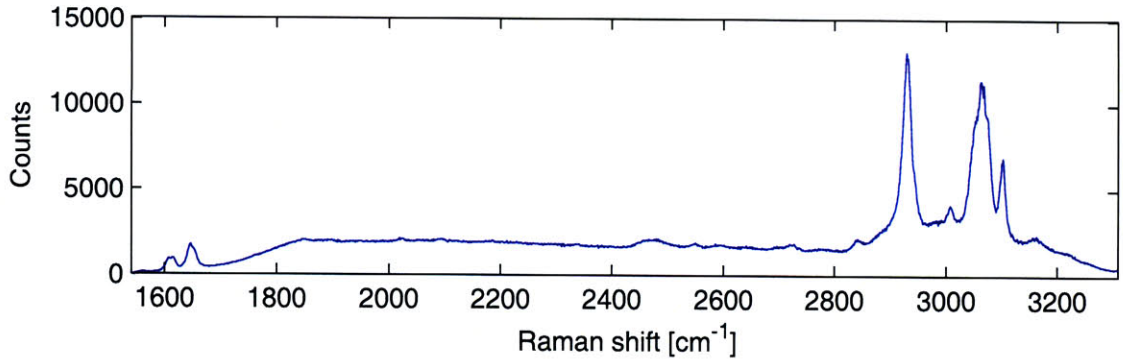


(b) High wavenumber background spectrm (80 mW, 60 frames @ 0.25 s/frame)

Figure 4-4: Raman probe background signal.



(a) Fingerprint (100 mW, 60 frames @ 0.25 s/frame)



(b) High wavenumber (80 mW, 60 frames @ 0.25 s/frame)

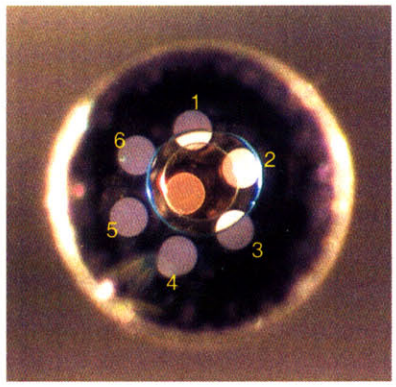
Figure 4-5: Raman probe acetaminophen spectrum.

in the $300 - 600$ cm^{-1} range.

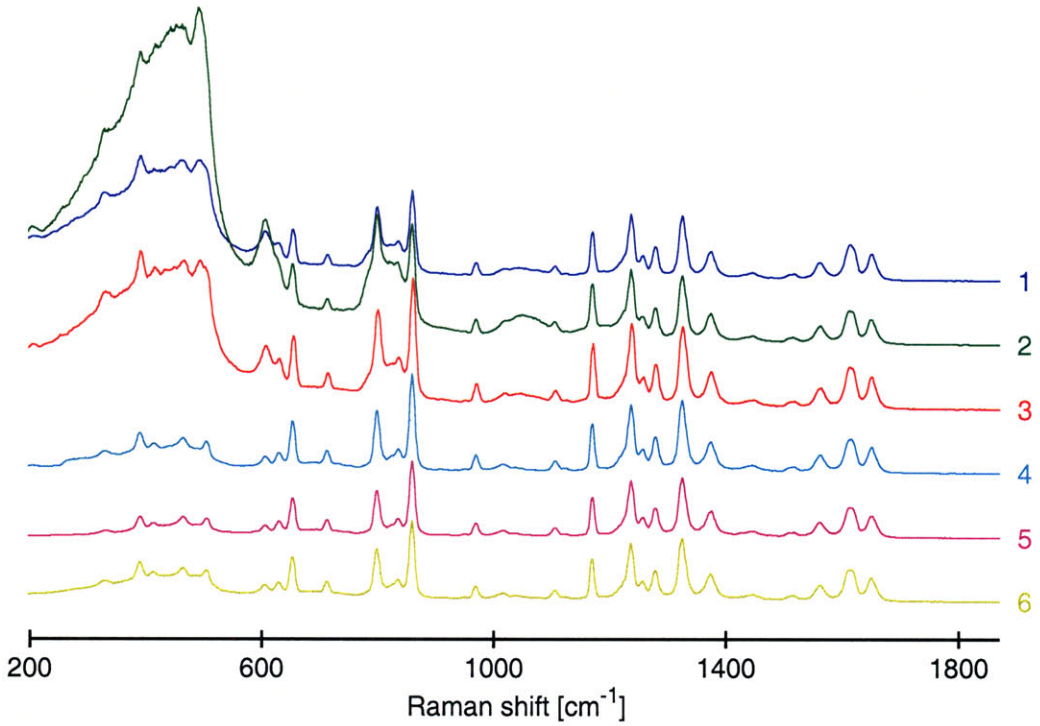
However, Figure 4-6 shows an example of a poorly fabricated Raman probe. In this case, the fiber bundle was misregistered with the patterned filter, leaving several of the collection fibers uncovered by the longpass filter (Figure 4-6(a)). This probe was used to acquire the Raman spectrum of acetaminophen powder (Figure 4-6(b)). In the improperly filtered fibers (1, 2, and 3), Raman bands from the silica are clearly visible at $300 - 530$ cm^{-1}, at 603 cm^{-1}, and at 809 cm^{-1}. While not entirely absent, the silica Raman signal is smaller in fibers 4 and 5, which are properly filtered.

4.2 Raman probe performance with distance

We expect the intracoronary Raman catheter will follow the same imaging paradigm as intravascular ultrasound, in which a small diameter catheter is placed within the artery. The catheter is free to move laterally within the lumen, and the distance between the



(a) Photograph



(b) Spectrum of acetaminophen powder

Figure 4-6: An improperly filtered Raman probe.

catheter and the arterial wall can vary with time. Because the distance between the Raman probe optics and the artery directly affects the total Raman signal collection, we explore the relationship between distance and signal intensity in saline and whole blood.

Using the forward-viewing filtered fiber bundle described in the previous section, we investigated Raman signal strength as a function of probe-tissue separation distance, with both saline and blood as intervening media. First, the probe was mounted onto a vertically oriented translation stage and placed in contact with *ex vivo* human aorta ($z = 0$ mm). A contact Raman spectrum was acquired. The probe was then translated away from the tissue at fixed steps ($\Delta z = 250$ μm), and Raman spectra were acquired at each step.

With saline as the intervening medium, the magnitude of the Raman spectrum decreases as the distance between the probe and the tissue increases, but the spectral features remain the same (Figure 4-7(a)). The Raman signal strength was quantified by integrating the spectrum over an arbitrarily chosen band ($2910 - 2940$ cm^{-1}). The Raman signal strength decreases to 50% at a distance of approximately 2 mm (Figure 4-7(b)).

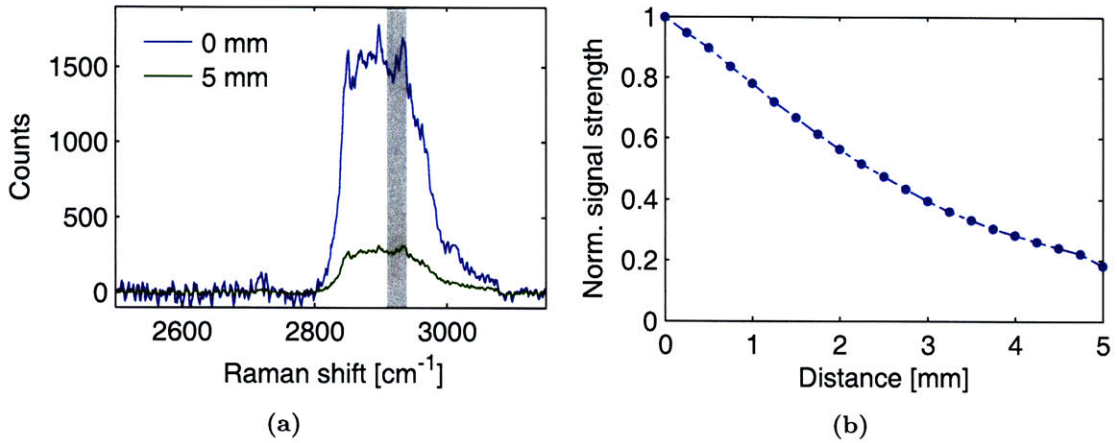


Figure 4-7: Raman probe performance through varying distance of saline. (a) High wavenumber Raman spectra of *ex vivo* human aorta in contact (blue) and through 5 mm of saline (green). The same spectral features are present in both spectra. To evaluate signal strength, spectral intensities were integrated in the region between the dashed lines. (b) Raman signal strength as a function of distance through saline.

The same experiment was repeated with porcine blood as the intervening medium between the probe and the tissue. As the distance through blood increases, the overall Raman signal intensity decreases. Furthermore, the collected Raman spectrum transitions from that of the tissue in contact to that of blood alone (Figure 4-8(a)). For distances greater than 2 mm, the Raman probe measures the Raman spectrum of the blood only, with no

signal being collected from the tissue. Thus the relevant parameter in this case is not the total signal strength, but the strength of the tissue Raman signal, as opposed to the blood Raman signal. This was evaluated by fitting each collected Raman spectrum as a linear combination of two basis spectra: the tissue contact spectrum ($z = 0$ mm) and the blood only spectrum ($z = 3$ mm). Examining the normalized tissue fit coefficient as a function of distance (Figure 4-8(b)), the tissue Raman signal strength decays to 50% at $z = 0.5$ mm. For $z > 1$ mm, the collected Raman spectrum is dominated by blood. Blood is a non-negligible absorber in the near infrared region [114], so as the probe-tissue separation distance increases, both the excitation light reaching the tissue and the returning Raman light are attenuated.

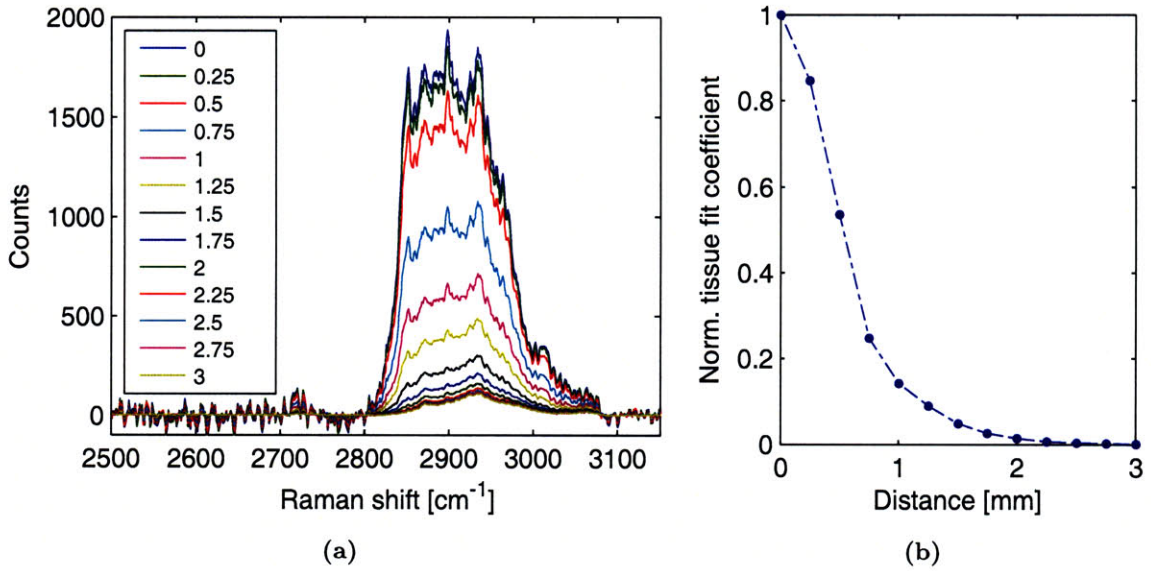


Figure 4-8: Raman probe performance through varying distance of blood. (a) High wavenumber Raman spectra of human aorta through varying distances of blood. As the distance through blood increases, the spectrum morphs from the pure tissue spectrum ($z = 0$ mm) to the pure blood spectrum ($z = 3$ mm). (b) Tissue fit coefficient as a function of distance. For probe-tissue distance greater than 1 mm, the measured Raman spectrum is predominantly that of blood.

These experiments demonstrate that it is not feasible to conduct Raman spectroscopy through more than ~500 μm of blood using this probe. Thus intracoronary Raman spectroscopy requires either a saline flush to remove blood from the artery before imaging, or a catheter that is designed to maintain contact with the arterial wall.

4.3 Discussion

One of the challenges that has limited the development of small diameter Raman probes is the need for distinct optical filters for the excitation and collection channels. Previous researchers have utilized micro-optical machining techniques to create separate filters, which were then incorporated into Raman probes. However, these filters and the resulting Raman probes were difficult and time-consuming to fabricate. In this work, we have utilized custom-made monolithic patterned filters, which greatly simplifies the Raman probe fabrication steps. Using a self-assembling fiber bundle and a patterned monolithic filter module, our Raman probe is 775 μm in diameter, though it can be re-rengineered to be as small as 500 μm without losing signal collection area.

Since the Raman probe will be incorporated into a catheter that can move freely within the arterial lumen, it is important to understand how the signal changes as a function of distance to the arterial wall. We experimentally investigated this, using both saline and porcine blood as the intervening medium. In saline, as the probe-tissue distance increases, the magnitude of the collected Raman signal decreases. The spectral signature of the acquired Raman spectrum remains constant for all distances.

However, in blood, the collected Raman spectrum has contributions from both the tissue and the blood, and for distances greater than ~1 mm, the Raman probe collects only the blood Raman spectrum. This is explained by the fact that the excitation light is attenuated before it reaches the tissue, resulting in a small tissue Raman signal that is also attenuated as it propagates to the collection fibers. Thus performing intracoronary Raman spectroscopy through blood over distances greater than ~0.5 mm is not feasible. There are two possible solutions; one is to design a catheter than can maintain contact with the arterial wall. The other, clinically preferred approach is to use a saline purge during imaging to evacuate the blood.

Using a forward-viewing contact Raman probe, Motz *et al.* have demonstrated that incorporating a lens at the distal end of the filtered fiber bundle can increase the magnitude of the collected Raman signal [86]. However, an intracoronary catheter may have variable probe-tissue separation distance, and the effectiveness of a distal lens is likely to vary with the distance between the probe and the tissue surface. Experimentally exploring the lens parameter space is cumbersome and cost-ineffective, so in Chapter 6 we develop a simulation

technique that can be used to computationally simulate Raman probe performance and thus aid in comparing probe designs.

4.4 Conclusion

In this chapter, we have demonstrated a simple Raman probe incorporating a self-assembling fiber bundle and shortpass and longpass filters patterned on a monolithic substrate. Ultimately the Raman probe will be incorporated into a catheter that will be able to freely move within the arterial lumen; thus the probe will be at variable distances from the arterial wall. Thus we investigated the probe's signal collection as a function of probe-tissue separation distance, through saline and blood. With blood as the intervening medium, the collected Raman spectrum is dominated by blood after a distance of $\sim$1 mm, suggesting a fundamental limitation that precludes conducting intracoronary Raman spectroscopy through blood. In saline, as the distance between the probe and the tissue surface increases, the Raman signal magnitude decreases, and spectral features are retained.

Chapter 5

Raman Catheter in a Xenograft Model

In this chapter we describe the conversion of the Raman probe in Chapter 4 into a side-viewing Raman catheter. The performance of the prototype Raman catheter was tested *in vivo* in a human-swine xenograft model, in which human coronary arteries obtained from cadaver are grafted onto a living swine.

5.1 Raman catheter

The Raman catheter is based on the forward-viewing Raman probe described in Chapter 4, consisting of a 6-around-1 fiber bundle bonded to a patterned filter module. The fiber bundle was converted into a side-viewing Raman catheter by incorporating a 45° mirror and placing the entire assembly into an appropriate outer sheath, using off-the-shelf components (Figure 5-1).

To hold the bundle and mirror in place, a window was cut in a 1.5 mm outer diameter stainless steel tube. The filtered fiber bundle and a 45° aluminum-coated rod mirror (NT47-628, Edmund Optics; Barrington, NJ) were epoxied into the tube, taking care not to allow any epoxy to leak between the filter and the mirror. A 1.5 m nylon sheath was bonded to the proximal end of the steel tube to house the loose optical fibers. A second, 2 cm nylon sheath was bonded to the distal end of the tube and modified to provide a rapid-exchange guide wire port. Clear medical grade heat shrink tubing was used to seal the assembly. When the catheter was held in contact with a sample, the laser spot size was approximately 1 mm

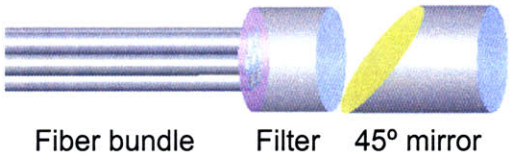
in diameter. Estimated from Monte Carlo simulations, we believe the catheter's sampling volume in arterial tissue was approximately 1 mm^3, when the catheter was in contact with the arterial surface.

The materials and geometry of the side-viewing catheter lead to a more intense and more feature-rich background signal than the forward-viewing filtered fiber bundle (Figure 5-2). The FP background spectrum is dominated by silica bands, but it also has strong bands at 855, 1288, 1614, and 1725 cm^{-1}, arising from the heat shrink tubing. This is one of the challenges for FP Raman probes; despite the fact that the heat shrink tubing thickness is only $\sim$ 25 μm, it still contributes appreciably to the catheter background spectrum in the region of interest. The heat shrink also contributes a band in the HW region at 3083 cm^{-1}, but this is fortunately outside the Raman shift region of interest. The catheter background spectrum is much larger in magnitude than that of the forward-viewing Raman probe. This is probably due to two geometry related effects: 1) the addition of the mirror and stainless steel tube may cause strong reflections to go back into the catheter and increase the background, and 2) there may be variability in the catheter fabrication process.

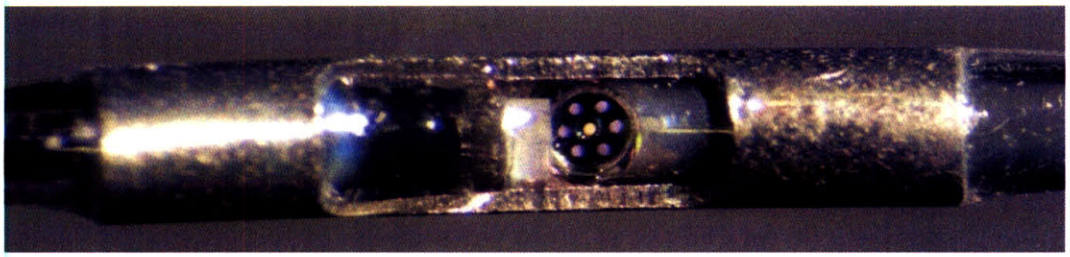
5.2 Xenograft procedure

5.2.1 The xenograft model

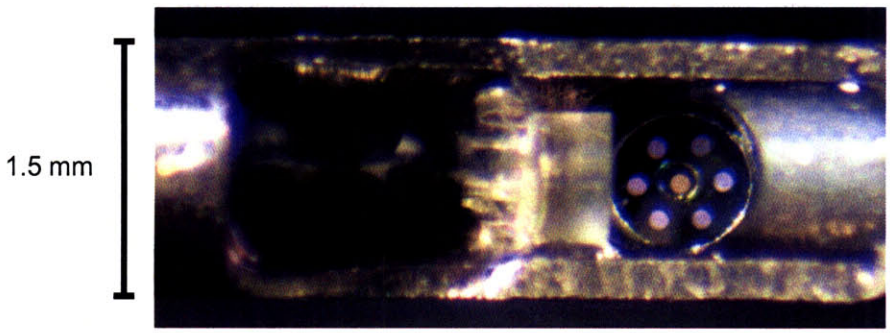
The prototype Raman catheter was tested in a human-swine xenograft model, in which a diseased human coronary is harvested from a cadaver, sutured on top of a living swine's heart to simulate cardiac motion, and connected to the swine's arterial circulation to provide physiologic blood flow conditions [115]. Following grafting, an intracoronary device can then be advanced through the grafted human artery. The human-swine xenograft accurately models the blood flow, motion, and compression and expansion of native coronary arteries, which enables testing of intracoronary devices for diagnosing real human pathology in an environment closely resembling human coronary physiology.



(a) A schematic of the optics



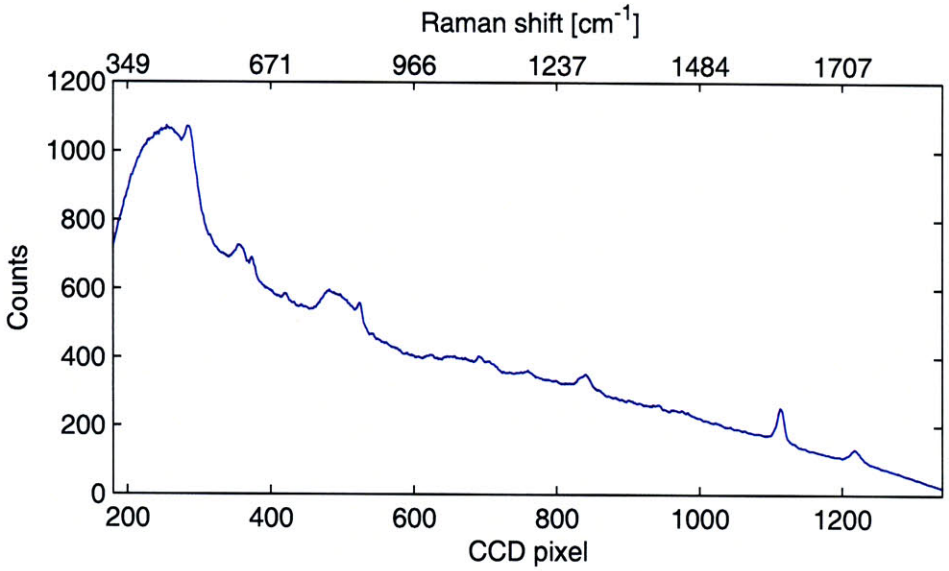
(b) Side-viewing catheter



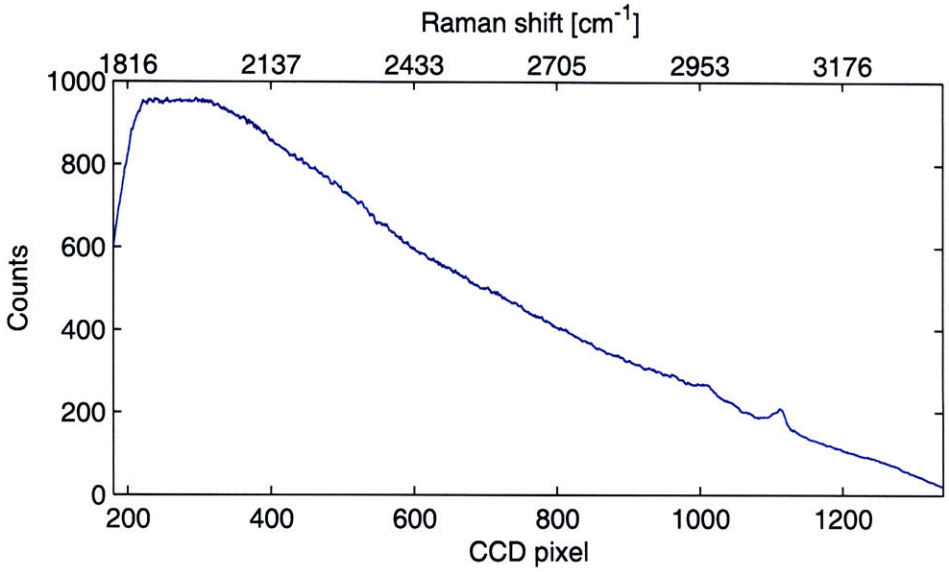
(c) Close-up view

(d) Distal tip of the catheter

Figure 5-1: First generation Raman catheter. The Raman catheter consists of a filtered fiber bundle and a 45° mirror enclosed in an outer sheath. The filtered fiber bundle and the mirror were epoxied into a steel tube (1.5 mm diameter). The image of the illuminated fibers at the fiber/filter interface is reflected by the mirror.



(a) FP Raman catheter background



(b) HW Raman catheter background

Figure 5-2: Raman catheter backgrounds as measured in air. The Raman catheter has background signals arising from silica and the other catheter materials.

Surgical procedure

Coronary arteries were harvested from an adult cadaver heart.[1] Segments of the coronary artery were prosected *en bloc* with surrounding muscle and epicardial fat to preserve architecture. The major side branches were ligated to prevent leakage and standard Luer lock connectors were attached to each end of the arterial segment. The grafts were preserved in cold saline until implantation.

At the start of the procedure, a female Yorkshire swine weighing approximately 40 kg was anesthetized, intubated, and ventilated with maintenance anesthesia of isofluorane. Heart rate and rhythm, temperature, and oxygen saturation were continuously monitored throughout the procedure. In accordance with the standard of care for catheterization procedures, heparin was administered to the swine to prevent the formation of blood clots.

The animal's heart was surgically exposed, and the proximal end of the cadaver artery was grafted to the animal's carotid artery via a cannula. The distal end of the cadaver artery was then grafted to the animal's femoral artery through another cannula. This procedure established blood flow from the swine's carotid artery, through the first cannula, and into the grafted artery. Blood was then returned to the swine's circulation via the second cannula to the femoral artery.

After the graft was in place, the side-viewing Raman catheter was introduced through the proximal cannula and spectra were taken at two distinct sites of the grafted artery. At the end of the experiment, the animal was euthanized. The human coronary artery segment was harvested, fixed in formalin, and sent for histological processing.

5.2.2 Data acquisition

Raman spectra were acquired from the grafted human artery *in vivo*, using illumination powers of 100 mW at 830 nm and 85 mW at 740 nm. FP and HW spectra were acquired over 40 frames at 0.25 s each for a total acquisition time of 10 s at each site, enabling us to evaluate spectral quality as a function of integration time. Based on previously reported results [83, 87] and *ex vivo* experiments performed in our laboratory [116], we believe that these laser fluences do not cause arterial tissue damage or blood coagulation. After each Raman measurement, the interrogated site (identified by visualizing laser light

[1]The xenograft procedure was performed at the Tufts-New England Medical Center, under approval from the Institutional Animal Care and Use Committee (Protocol #55-06).

transmitted through the artery wall) was marked with green ink for subsequent correlation with histology.

5.3 Results

Raman spectra were extracted from the raw spectra using established post-processing methods. First, a specified number of raw data frames were averaged together and divided by the white light reference spectrum. The tissue fluorescence spectrum was removed by subtracting a fifth (FP) or sixth (HW) order polynomial, and the catheter background was removed by subtracting a processed background spectrum, scaled to remove the bands resulting from the heat shrink tubing (1614 cm^{-1} and 3083 cm^{-1}). To determine the minimum necessary acquisition time, we averaged various numbers of frames together and analyzed the resulting processed Raman spectra. The Raman spectra shown in Figure 5-3 correspond to an integration time of 4 s (16 frames).

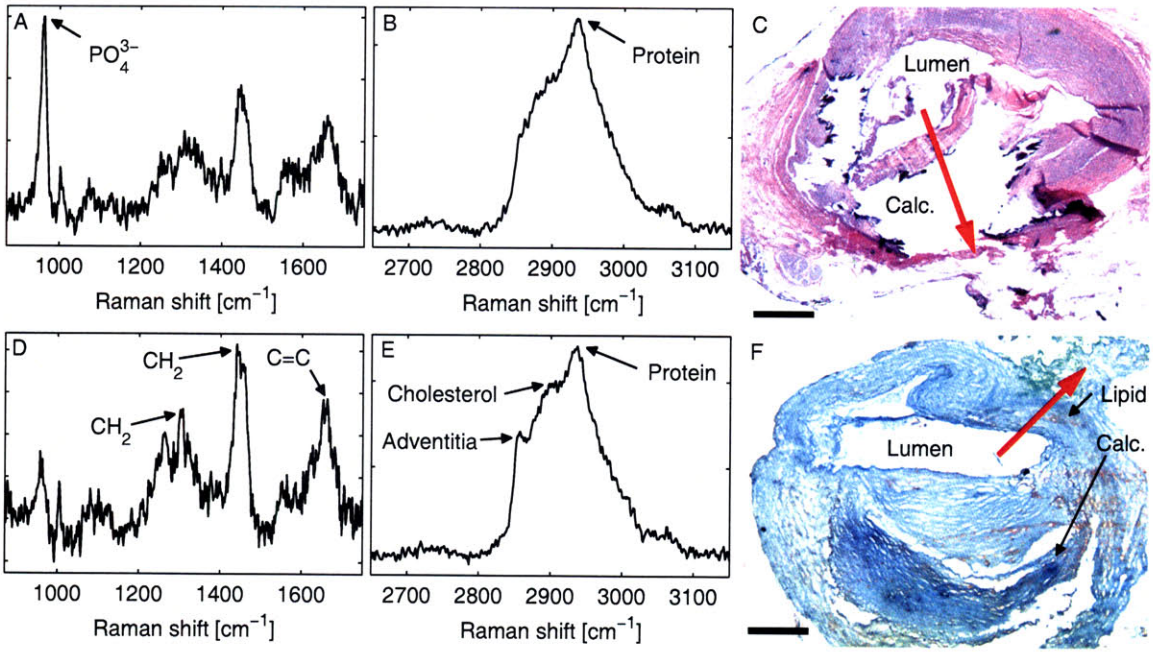


Figure 5-3: Raman spectra obtained in a human artery grafted to the beating heart of a living swine. Spectra were processed from 16 averaged frames, representing a total acquisition time of 4 s. (a) FP and (b) HW spectra obtained at site 1, with (c) corresponding H&E histology. (d) FP and (e) HW spectra obtained at site 2, with (f) corresponding Oil-red-O histology. Black arrows denote brown-orange staining of lipid in the direction of the spectral measurement and dark blue staining of a large calcific nodule nearby. Red arrows in (c) and (f) denote spectral measurement locations and directions. Calc. is calcium hydroxyapatite. Scale bars, 0.5 mm.

Raman spectra from the two sites differ significantly. In site 1 (Figures 5-3(a) and 5-3(b)), the FP Raman spectrum contains a prominent peak at 960 cm^{-1} (corresponding to the phosphate stretch of calcium hydroxyapatite) and minimal lipid. The Raman spectrum is consistent with histology, which demonstrates a large calcific nodule (Figure 5-3(c)). Spectra obtained at the second site (Figures 5-3(d) and 5-3(e)) show peaks representative of adventitial triglycerides (CH_2 bends at 1301 and 1440 cm^{-1} and C=C stretch at 1654 cm^{-1} in the FP spectrum and ~2836 cm^{-1} in the HW spectrum). There is an increased contribution from cholesterol (~2870 cm^{-1}) in the HW spectrum, which may correspond to a small amount of lipid staining seen on the Oil-red-O histology section (Figure 5-3(f)). A minute calcium hydroxyapatite signal is also present, which may result from the large, adjacent calcific nodule.

5.4 Discussion

In this chapter, we have demonstrated a catheterized version of the Raman probe described in Chapter 4, capable of acquiring both FP and HW Raman spectra *in vivo*. This catheter exhibits a higher magnitude, more feature-rich background spectrum than the forward-viewing Raman probe, arising from the extra materials associated with the catheter. In particular, in the FP region, the heat shrink tubing used to seal the catheter optics has a Raman spectrum with sharp peaks that overlap with the tissue Raman signal. The heat shrink tubing's added band in the HW region does not overlap with the tissue Raman signal. While the catheter background can be mitigated by carefully choosing catheter materials, the relative paucity of bands in the HW region highlights a strong advantage of the HW region over the FP region for conducting catheter-based Raman spectroscopy.

The present catheter has a relatively large outer diameter of 1.5 mm (4.5 Fr), due to the use of the 1 mm diameter off-the-shelf mirror. The catheter's size-limiting factor is the 775 μm diameter of the fiber bundle and filter. Thus the overall catheter diameter can be reduced to less than 1 mm by simply replacing the side-viewing mirror and outer sheath, which will bring the catheter diameter on par with that of commercially available intracoronary imaging catheters.

The catheter was tested *in vivo* in a xenograft procedure, a model developed for testing intracoronary devices on human pathology in an environment that is physiologically similar

to humans. Spectral differences were seen between a calcified site and a lipid-rich site in both the FP and HW regions.

The present catheter has not been optimized for optical design, and thus it operates in a relatively low SNR regime. The xenograft results demonstrate several advantages to the HW region. First, the signal in the HW region is ~6 times higher than in the FP region, resulting in higher SNR. Second, as mentioned, catheter design in the FP region is more difficult because many polymer materials that might be incorporated into the catheter contribute a large Raman and/or fluorescent background signal. Fewer materials have a Raman signal that overlaps with the HW region of interest, so HW catheter design may be simpler. Indeed, a HW catheter may not need optical filters at all [88]. Finally, the lack of a strong background signal in the HW region also simplifies the data processing, since subtraction of an overlapping background spectrum is not required. These advantages suggest that the HW region is preferred over the FP region for catheter-based Raman spectroscopy, but these considerations must be balanced against the lower information content available in the HW region.

The SNR of the collected Raman spectra can be improved by optimizing the catheter. Because of the side-viewing geometry, even when the catheter is held in contact with the tissue, the probe-tissue separation distance, as defined in Chapter 4, is approximately 1 mm in air. Thus the signal intensity for the Raman catheter is relatively low compared to the forward-viewing Raman probe. The catheter optical geometry can be optimized significantly. For example, simply decreasing the outer diameter of the catheter would decrease the probe-tissue separation distance and result in increased signal. The 45° aluminum-coated mirror could be replaced with a gold or silver mirror to increase reflectivity. Furthermore, the 45° mirror and resulting air gap could be entirely replaced by a mirror-coated prism, which would be better index-matched to the filter and thus maintain a smaller excitation spot size. The optical simulation technique in Chapter 6 may be a useful tool for exploring these and other alternative designs.

5.5 Conclusions

In this chapter, we have converted the foward-viewing Raman probe into a side-viewing Raman catheter and demonstrated it in *in vivo* in a xenograft model. Spectral differences

were seen between the two interrogated sites in the grafted artery. For catheter-based Raman spectroscopy, the HW region has several advantages over the FP region, including potentially simpler probe design, simpler data processing, and higher overall signal. Future work should concentrate on increasing the Raman signal SNR through optimization of the catheter optics.

Chapter 6

Optical Modeling for Catheter Design

The success of an intracoronary Raman catheter depends on its ability to acquire spectral data with sufficiently high signal-to-noise ratio for accurate spectral analysis. Furthermore, the SNR must remain high throughout the catheter's expected operating conditions, *i.e.*, throughout the range of expected catheter-tissue distances. While we have demonstrated a catheter without a lens in Chapters 4 and 5, questions remain about whether or not a lens at the distal tip would increase Raman collection efficiency. Since experimental optimization of the probe design is costly and time-consuming, we have chosen to develop a model-based approach that in the future will enable optimization of Raman catheter designs.

In this chapter, we describe a optical simulation technique that enables assessment of sampling volume and collection efficiency for a given Raman probe design, prior to fabricating and experimentally characterizing the probe. This optical simulation technique accounts for 1) the propagation of excitation laser light through the excitation fiber and into the sample, 2) the resulting generation of Raman light within the sample, and 3) the subsequent propagation of Raman scattered light through the sample and into the probe's collection optics. The simulation technique combines several existing software tools, namely a freely-available Monte Carlo software package for propagation through tissue, and a commercial package for modeling light propagation through the probe's optical components. In this chapter we describe the simulation technique and present experimental validation results.

6.1 Monte Carlo Modeling

Light propagation in biological tissue is a diffusive process in which each photon takes a random walk through the tissue. For turbid media, analytical solutions to the radiative transport equation exist only for simple geometries, so Monte Carlo simulations are often used as a numerical tool to simulate light transport in tissue. In this section, we provide an overview of the Monte Carlo model, a mathematical basis for incorporating Raman scattering into the Monte Carlo model, and a convolution method for reducing computation time.

6.1.1 Standard Monte Carlo Model

Monte Carlo simulations are commonly used to model light propagation in tissue [117]. In Monte Carlo simulations, light propagation is modeled by stochastically tracing the propagation paths of a large number of photons (typically $> 10^5$), to converge on the real profile of interest. The basic procedure for a fixed weight, variable step-size Monte Carlo simulation is shown in Figure 6-1. First, a photon is initialized with its position and propagation direction determined by the desired simulation geometry. At each step, the photon is moved a random distance Δs in the tissue, where it can be scattered, absorbed, internally reflected, or transmitted out of the tissue. In a fixed-weight Monte Carlo model, when a photon is absorbed, propagation stops, and the photon's absorption position is recorded. If a photon is elastically scattered or internally reflected, a new step size and random propagation direction are determined, and the photon continues to propagate until it is either absorbed or remitted from the tissue. This basic process is repeated until the desired number of photons have been propagated. The optical properties of the tissue determine the propagation distance and angle at each step, as described in further detail below.

The simulation geometry is typically defined with the z axis positive going into the tissue (Figure 6-2(a)). Each photon is described by its position (x, y, z), its propagation direction, expressed by direction cosines (μ_x, μ_y, μ_z), and its weight, w. The direction cosines are

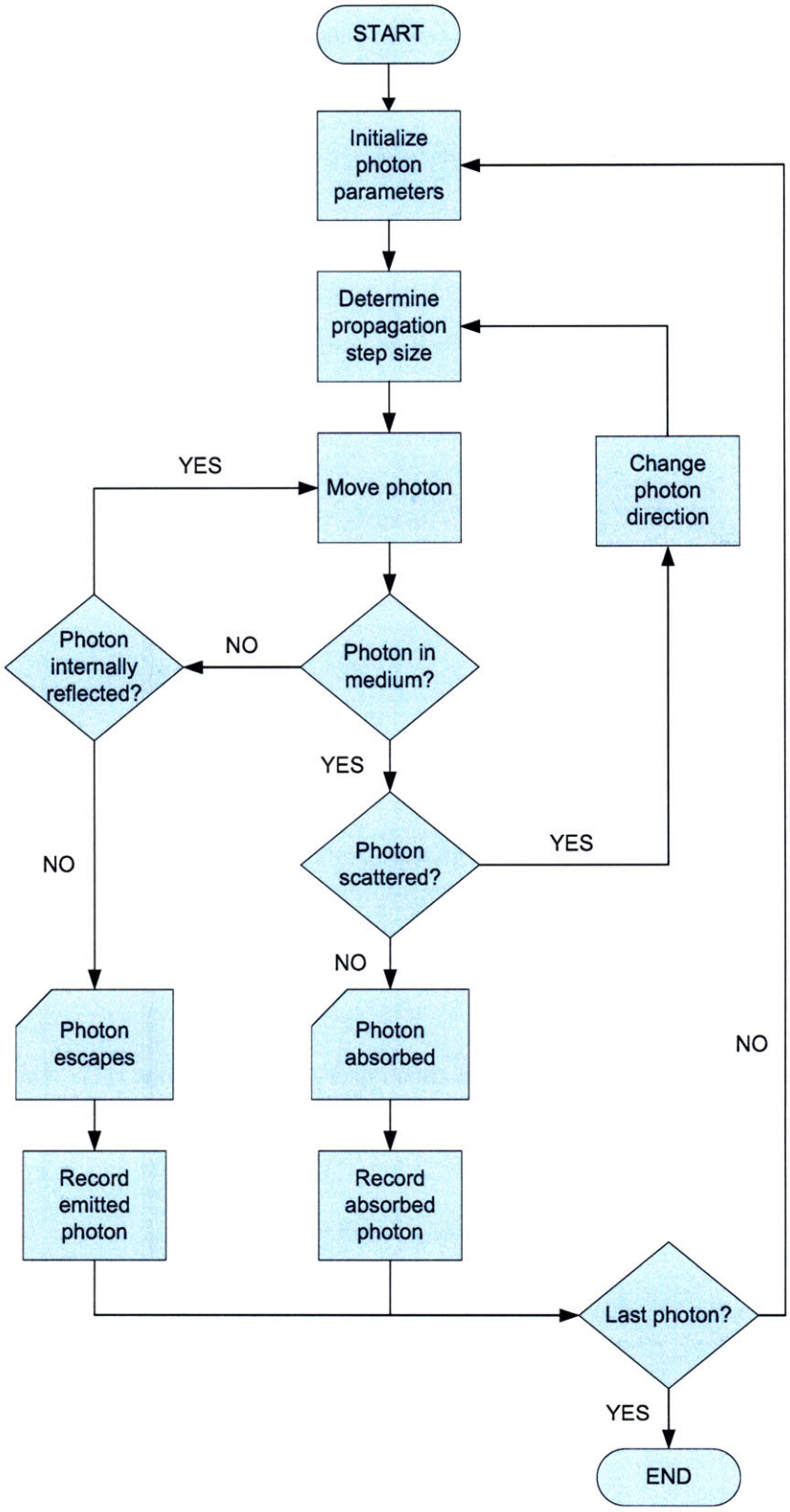
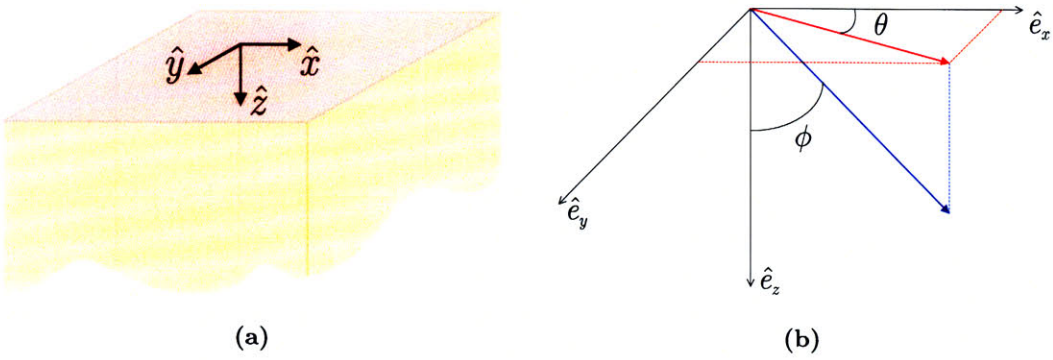


Figure 6-1: Block diagram of the basic Monte Carlo simulation method.

related to propagation angle (Figure 6-2(b)) by

$$\begin{aligned} \mu_x &= \sin(\phi)\cos(\theta) \\ \mu_y &= \sin(\phi)\sin(\theta) \\ \mu_z &= \cos(\phi). \end{aligned} \tag{6.1}$$

For example, in a collimated pencil beam, the standard for Monte Carlo simulations, all photons are launched from the same initial position $(0, 0, 0)$, with all photons directed vertically downward into the tissue ($\mu_x = 0, \mu_y = 0, \mu_z = 1$).



(a) **(b)**

Figure 6-2: Monte Carlo simulation geometry. (a) Coordinate axes for tissue, with positive z indicating depth. (b) A photon's propagation direction is defined by ϕ, its deflection angle from a reference axis, and θ, its azimuthal angle about that axis.

In the photon model of propagation through a turbid medium [117], the probability of a photon propagating through a distance l of tissue without being absorbed is modeled as an exponential process:

$$p_a(l) = \mu_a \exp(-\mu_a l), \tag{6.2}$$

where μ_a is the absorption coefficient. Similarly the probability of the photon propagating through a distance l before being scattered is

$$p_s(l) = \mu_s \exp(-\mu_s l), \tag{6.3}$$

where μ_s is the scattering coefficient. In the standard Monte Carlo method, the two exponential processes describing absorption and scattering are merged to determine the proba-

bility of a photon traveling a distance l with no interaction:

$$p_t(l) = \mu_t \exp(-\mu_t l), \tag{6.4}$$

with

$$\mu_t = \mu_a + \mu_s. \tag{6.5}$$

This probability density function (pdf) is used to generate the step size, $\Delta s = l$, for each iteration of the Monte Carlo simulation; the step size is interpreted as the distance the photon travels before a scattering or absorption event occurs. After the photon has been propagated through the distance Δs, the exponential pdf $p_t(l)$ is split to yield the probability that the photon is absorbed,

$$\Pr(\text{photon is absorbed}|\text{an event occurred}) = \frac{\mu_a}{\mu_t}, \tag{6.6}$$

and the probability that the photon is scattered,

$$\Pr(\text{photon is scattered}|\text{an event occurred}) = \frac{\mu_s}{\mu_t} = \text{albedo}. \tag{6.7}$$

To decide whether the photon is absorbed or scattered, a uniformly distributed random number, ξ, is generated, and if $\xi < \frac{\mu_a}{\mu_t}$, the photon is absorbed, otherwise it is scattered.

For scattered photons, a new propagation direction is determined, requiring two new angles. The deflection angle, ϕ, defined with respect to the photon's current propagation direction, is stochastically determined by a scattering phase function, commonly the Henyey-Greenstein phase function [118],

$$\Pr(\cos\phi) = \frac{1 - g^2}{2(1 + g^2 - 2g\cos\phi)^{3/2}}, \text{with } \phi \in (0, \pi), \tag{6.8}$$

where g is the anisotropy parameter, defined as $g = E[\cos\phi]$. The azimuthal angle is taken to be uniform, $\theta \in [0, 2\pi]$.

In practice, the fixed weight Monte Carlo is rarely used, due to the long computation required to reach convergence. Instead, a weighted Monte Carlo model is used, where instead of propagating individual photons, photon packets, or groups of photons that travel together, are propagated. At each step, a portion of the photons are absorbed $\left(\frac{\mu_a}{\mu_t}\right)$, so the

corresponding weight is dropped from the photon packet. The remaining photons in the packet continue to propagate as usual.

Monte Carlo simulation requires the propagation of a large number of photons to ensure statistical convergence of the results. The required number of photons depends on the geometry and the optical properties used in the simulation, but is typically on the order of 10^5 photons for biological tissue.

Typical outputs of a standard Monte Carlo simulation include the 3D absorption profile in the tissue, $A(x, y, z)$, and the 2D distribution of photons remitted from the surfaces of the tissue $E(x', y', \phi, \theta)$, *i.e.*, remitted flux and exit angle as a function of x' and y' on the tissue surface.

6.1.2 Incorporating Raman scattering into the Monte Carlo model

Previous researchers have added fluorescence to the Monte Carlo model by assuming that the absorbed photons will either be converted to heat or re-emitted as a fluorescent photon [119]. Similarly, Raman scattering may be modeled by assuming that all scattered photons will either be elastically scattered or inelastically scattered, the latter generating a Raman photon [120]. Each Raman photon is assumed to be isotropically generated, *i.e.*, the photon's initial phase and direction are uniformly distributed.

Raman scattering can be considered analogously to Eq. (6.2) and (6.3). The probability of a photon traveling through a distance l before being Raman scattered can be expressed as

$$p_R(l) = \mu_R \exp(-\mu_R l), \tag{6.9}$$

where μ_R is the Raman scattering coefficient. Absorption, scattering, and Raman scattering can again be merged to determine the probability of a photon traveling a distance l with no interaction:

$$p_{t'}(l) = \mu_{t'} \exp(-\mu_{t'} l), \tag{6.10}$$

with

$$\mu_{t'} = \mu_R + \mu_a + \mu_s, \tag{6.11}$$

where $\mu_{t'}$ now includes Raman scattering.

Directly using this model to simulate Raman scattering requires the propagation of

a very large number of photons ($> 10^{14}$ photons), because only a small percentage ($\sim$1 in 10^{10}) of all incident photons are Raman scattered. This simulation is computationally intensive.

Alternatively, since both absorption and Raman scattering are modeled as events that "destroy" the propagating photon, the two processes can be merged into a "destruction" event, by lumping the two parameters:

$$\mu_D = \mu_a + \mu_R. \tag{6.12}$$

Equation (6.11) can be rewritten as

$$\mu_{t'} = \mu_D + \mu_s. \tag{6.13}$$

Since Raman scattering is a very low probability event, $\mu_R \ll \mu_a$, and thus $\mu_{t'} \approx \mu_t$ and $\mu_D \approx \mu_a$. Thus, Raman scattering can be simulated using the standard Monte Carlo model, where photon absorption is now treated as photon destruction. After propagating a photon using the standard model, the photon destruction process can then be split into absorption and Raman scattering events, with

$$\Pr(\text{photon is absorbed}|\text{destruction event occurred}) = \frac{\mu_a}{\mu_D} = \frac{\mu_a}{\mu_a + \mu_R} \tag{6.14}$$

and

$$\Pr(\text{photon is Raman scattered}|\text{destruction event occurred}) = \frac{\mu_R}{\mu_D} = \frac{\mu_R}{\mu_a + \mu_R}. \tag{6.15}$$

The spatial distribution of Raman scattering events differs from the spatial distribution of absorption events by a constant factor, μ_R/μ_a. Thus, it is possible to efficiently simulate the generation of Raman scattered photons using the standard Monte Carlo model and scaling the resulting absorption profile appropriately.

One method for simulating the spatial distribution of remitted Raman scattered light might involve propagating a photon until it is absorbed by the medium, and then storing its absorption position as the generation position for a new Raman photon. Once the location of all Raman scattering events has been determined, the distribution of remitted Raman

photons can be evaluated by performing a separate, simplified Monte Carlo simulation, using optical property values appropriate for each Raman photon's wavelength. The simplified simulation utilizes the basic Monte Carlo model described in Section 6.1.1, in which only elastic scattering or absorption may occur. The initial position of each photon is the absorption location recorded above, and the new direction is isotropically generated (both ϕ and θ are uniform with $\phi \in [0, \pi]$ and $\theta \in [0, 2\pi]$). The weight of the photon packet remains unchanged. Throughout the simulation, the locations of all absorbed and remitted Raman photons are recorded.

This simulation is still computationally intensive. Fortunately, the Raman Monte Carlo model can be simplified to enhance computational efficiency by incorporating a convolution model, which has been previously demonstrated for simulations of fluorescent photons [121].

6.1.3 Convolution model for efficient Raman Monte Carlo simulations

For computational efficiency, a convolution model can be used to determine the absorption and emission profiles for the Raman scattered light. In this model, two effects are simulated separately: 1) the propagation of excitation photons and resulting isotropic emission of Raman photons, and 2) the propagation of the newly generated Raman photons. Results from each simulation step are then combined through convolution.

In the first step, a "Raman source," $E_{\text{Raman,gen}}(x, y, z)$, is simulated, representing the probability of a Raman photon being scattered at position (x, y, z) within the tissue. This source can be efficiently simulated by assuming that the spatial distribution of Raman scattered photons is a scaled version of the distribution of absorbed photons, as described in the previous section. A standard Monte Carlo simulation is run for the excitation wavelength, keeping track of absorbed photons, without propagating new Raman photons.

Next, we simulate the propagation of generated Raman photons. We consider a Raman point source, located at position (x, y, z). We assume that the remitted light distribution from each Raman point source is spatially invariant in x and y, *i.e.*, the distribution depends only on the depth, z, from which the photons originate and not the point source's (x, y) position. This assumption is valid when the tissue is modeled as a series of homogeneous semi-infinite layers. Thus for each depth z_0 of interest, one simulation is performed, consisting of an isotropically emitting point source located at $(x = 0, y = 0, z = z_0)$. This yields the Raman emission profile, $E_{\text{iso,em}}(x', y'; x = 0, y = 0, z = z_0)$, where x' and y' are

coordinates on the tissue surface, and $(x = 0, y = 0, z = z_0)$ is the location of the isotropic emitter.

Finally, for each depth of interest ($z_0 \in [z_{min}, z_{max}]$), the corresponding Raman source $E_{\text{Raman,gen}}(x, y, z = z_0)$, is convolved with $E_{\text{iso,em}}(x', y'; x = 0, y = 0, z = z_0)$ to determine the total remitted Raman light from that depth. To determine the total Raman emission profile, using superposition, the responses from each depth can be integrated:

$$E_{\text{Raman,emitted}}(x', y') = \int_{z_{min}}^{z_{max}} \left(E_{\text{Raman,gen}}(x, y, z_0) * E_{\text{iso,em}}\left(x', y'; 0, 0, z_0\right) \right) dz_0. \quad (6.16)$$

The output of this procedure, $E_{\text{Raman,emitted}}(x', y')$, represents the spatial distribution of the flux of the remitted Raman light. The distribution of exit angles of the remitted Raman photons can be calculated analogously.

6.2 Simulating Sampling Volume

Using the Raman Monte Carlo model described in Section 6.1, we can determine the Raman sampling volume of an optical system. Consider a probe in contact with tissue (Figure 6-3(a)), which may consist of a ring of collection fibers surrounding a central excitation fiber, as in Chapter 4. For a given optical geometry, we evaluate the percentage of $E_{\text{Raman,emitted}}(x', y'; x, y, z)$ that can propagate through the collection optics and ultimately be detected by the spectrometer's CCD.

Keeping track of this quantity on a voxel-by-voxel basis yields the sampling volume. We define a new quantity, $E_{\text{Raman,coll}}(x, y, z)$, as the percentage of $E_{\text{iso,em}}(x', y'; x, y, z_0)$ that is collected. Then we define

$$E_{\text{Raman,sampling volume}}(x, y, z) = E_{\text{Raman,gen}}(x, y, z) \times E_{\text{Raman,coll}}(x, y, z). \quad (6.17)$$

This quantity represents the total collected Raman signal originating from the voxel centered at (x, y, z), and it accounts for both the excitation geometry and collection geometry of the Raman probe. By summing and properly scaling these results, we can also define the Raman probe's collection efficiency.

6.2.1 Implementation

For practical reasons, we use two separate software packages to simulate the Raman probe sampling volume (Figure 6-3). For propagating photons through tissue, we use a freely available software package released by Lihong Wang and Steven Jacques, Monte Carlo for Multi-Layered media (MCML)[1] [122], which we have modified for this purpose. To propagate through the optical elements of the catheter, we use a commercially available ray tracing program, Zemax (Zemax Development Corporation; Bellevue, WA). Finally, we use MATLAB (The MathWorks; Natick, MA) to convert data between the preferred input/output formats for each software package and to compile the final results.

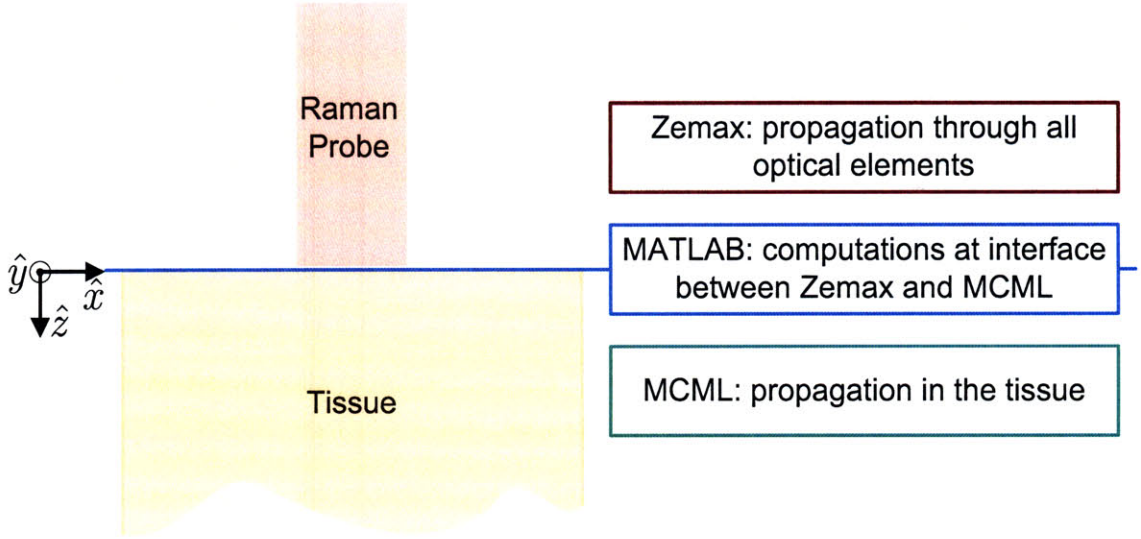
The general simulation procedure is as follows (Figure 6-3(b)):

1. **Propagate excitation light to the interface plane.** Zemax is used to propagate light from the laser source, through the excitation fiber and filter, through any distal optics in the probe, and finally to the interface plane in Figure 6-3. For all rays that hit the interface plane, the ray's position (x, y, z), direction cosines (μ_x, μ_y, μ_z), and weight, w, are recorded and stored in a file.

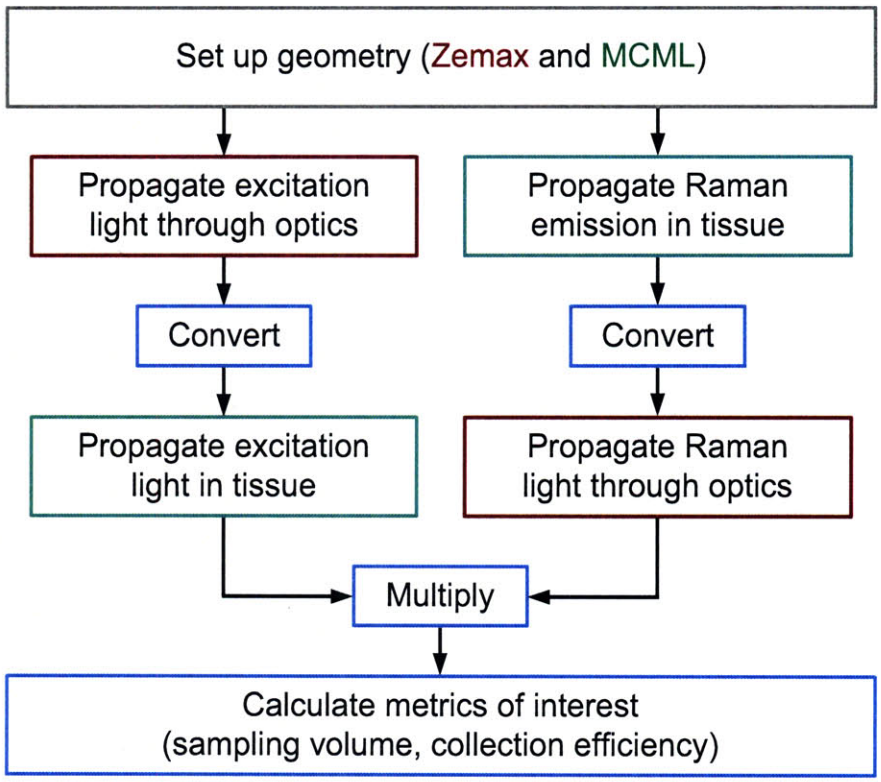
2. **Propagate excitation light from the interface plane and into the tissue to determine $E_{\mathbf{Raman,gen}}(x, y, z)$.** The stored excitation rays at the interface plane are then propagated through the tissue using our modified MCML. Absorption is recorded for each voxel, yielding the Raman source $E_{\text{Raman,gen}}(x, y, z)$.

3. **Propagate isotropically emitted Raman photons to the interface plane to determine $E_{\mathbf{iso,em}}(x', y'; 0, 0, z_0)$.** Independent of the first two steps, which are related to the excitation only, we perform the Raman emission simulations. For each depth of interest, we create a Raman point source, *i.e.*, we generate isotropic Raman photons in a voxel centered at $(x = 0, y = 0, z = z_0)$. The modified MCML is used to propagate the isotropically genererated photons through the sample, using optical properties appropriate for the Raman photons' wavelength. For all rays exiting the interface plane towards the catheter, the exit position (x', y'), direction cosines (μ_x, μ_y, μ_z), and weight, w is recorded. This yields $E_{\text{iso,em}}(x', y'; 0, 0, z_0)$.

[1]The source code and its extensive documentation can be downloaded from http://omlc.ogi.edu/software/mc/.



(a) Software packages



(b) Simulation procedure

Figure 6-3: Combined Monte Carlo/Zemax optical modeling. Zemax is used to propagate photons through the optical elements of the probe, and MCML is used to propagate photons through the tissue. MATLAB is used to convert inputs/outputs of both programs as necessary at the interface plane between the probe and the tissue.

4. **Propagate Raman photons from the interface plane and through the probe to determine $E_{\text{Raman,coll}}(x, y, z)$.** For each voxel in the sample, we wish to determine the percentage of generated Raman photons that can propagate through the probe to the spectrometer. To accomplish this, the recorded exiting Raman rays from step 3 are shifted by the center coordinates of the voxel, and Zemax is used to determine the appropriate percentage. In other words, to calculate $E_{\text{Raman,coll}}(x = x_1, y = y_1, z)$, we shift the distribution $E_{\text{iso,em}}(x', y'; 0, 0, z_0)$ by (x_1, y_1) to get $E_{\text{iso,em}}(x' + x_1, y' + y_1; x_1, y_1, z_0)$, and then propagate this distribution in Zemax. Thus, a separate Zemax calculation is performed for each voxel. The output of this step is $E_{\text{Raman,coll}}(x, y, z)$, a 3D matrix in which the value of each voxel represents the collection efficiency for Raman photons generated in that voxel.

5. **Multiply to determine the sampling volume, $E_{\text{Raman,sampling volume}}(x, y, z)$.** Finally, we use MATLAB to calculate the Raman sampling volume according to Eq. (6.17). This 3D matrix represents the relative Raman intensity that is collected from each voxel. The probe's collection efficiency can also be calculated from this matrix. It is important to note that the physical interpretation of the numerical values of $E_{\text{Raman,sampling volume}}(x, y, z)$ hinges upon accurate knowledge of μ_R, which is difficult to measure and generally not known for tissue. Nevertheless, these numerical values can be used to compare simulation results from one probe geometry to another and thus compare the relative performance of various probe designs.

6.2.2 Monte Carlo for Multi-Layered media (MCML)

By default, MCML is programmed to propagate a pencil beam into the tissue and to record absorption in the tissue, reflectance from the tissue, and transmission through the tissue, on a cylindrical grid. We have incorporated several modifications to MCML to make it suitable for modeling Raman sampling volume. First, we modified MCML to initialize photons with arbitrary initial position, direction, and weight. MCML was modified to read, from an accompanying input file, seven parameters for each ray: position (x, y, z), direction (μ_x, μ_y, μ_z), and photon weight, w. Second, MCML was modified to record the same seven parameters for each ray exiting the tissue; this output file is later converted to an input file for Zemax. Finally, we added a dynamically allocated 3D matrix to keep track of absorption

in the tissue as a function of (x, y, z).

6.2.3 Zemax

We used Zemax in non-sequential mode, which allows for ray tracing. To propagate the excitation light through the probe, a Zemax source is defined at the proximal end of the excitation fiber, to model the output of the laser source. A detector is placed at the distal surface of the probe (the interface plane) to record the rays that cross the interface plane; these rays are saved as a Zemax source file and later converted to an MCML input file.

To propagate the remitted Raman light from the interface plane and thus calculate Raman collection, a source file is defined at the interface plane, and detectors are defined at the proximal end of the collection fibers. The source file contains the MCML output rays from the isotropic emitter simulations. Rays from the source file are propagated back through the probe, and the total flux hitting each collection fiber detector is recorded as the collected Raman signal.

In the Zemax model, we only model the propagation through the first 4 mm of fiber, explicitly assuming that all rays that meet the fiber's propagation conditions will eventually propagate through the spectrometer to the CCD. Thus, our current model does not account for losses in the fiber, especially due to bending, coupling efficiency between the fiber and the spectrometer, losses in the spectrometer, or the quantum efficiency of the CCD.

6.2.4 MATLAB

MATLAB is used to perform computations at the interface plane between MCML and Zemax, mainly consisting of translating ray information from the modified MCML input/output file format to and from the Zemax input/output file format. The MATLAB wrapper code also converts the position units from MCML's cm units to Zemax's default mm units. Finally, MATLAB is used in step 5 to calculate the sampling volume and to visualize the results.

6.3 Validation

In this section we describe experimental validation of the probe sampling volume simulations. One method of measuring Raman probe sampling volume is to sweep a Raman point

source through the sampling volume and measure the resulting Raman signal as a function of position. In lieu of a true point source, we have utilized a small diamond chip immersed in a scattering phantom solution [94]. The diamond is ideal for this purpose because it has a strong and distinct Raman scattering signal (Figure 6-4). In this section, we describe the experiment and the details of the corresponding simulation.

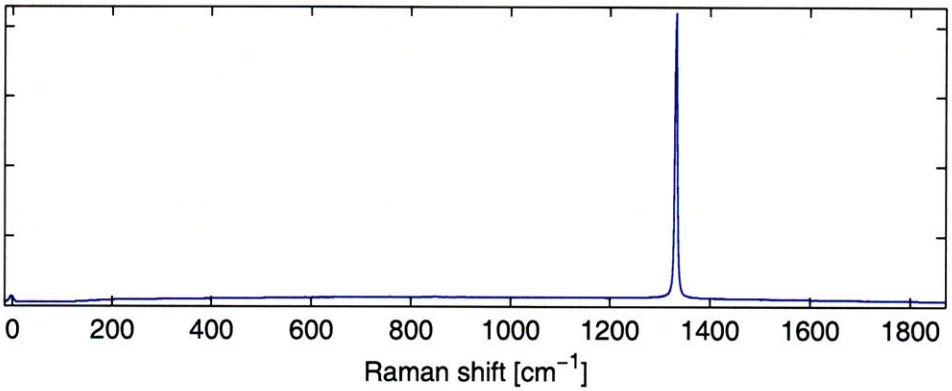


Figure 6-4: Raman spectrum of single crystal diamond, exhibiting a strong peak at 1332 cm^{-1}.

6.3.1 Experimental setup

The general experimental setup is shown in Figure 6-5. A small cylindrical container was modified to have a microscope cover glass as a side wall, and the container was filled with a scattering phantom solution with known optical properties. A dual fiber Raman probe, consisting of a single excitation fiber and a single collection fiber, was placed external to the container and against the cover glass. A small diamond chip, mounted on the tip of a hypodermic needle, was submerged in the solution. The diamond chip was scanned through the sampling volume of the probe, with Raman spectra acquired at every 20 μm within the volume.

Phantom

The phantom solution was an aqueous mixture of polystyrene microspheres and ink. The scattering coefficient was controlled via the concentration of 1 μm diameter polystyrene microspheres (Duke Scientific 5100A, Thermo Fisher Scientific; Waltham, MA). The absorption coefficient was modified by adding india ink to the solution.

The container holding the phantom solution was a modified centrifuge tube (Corning).

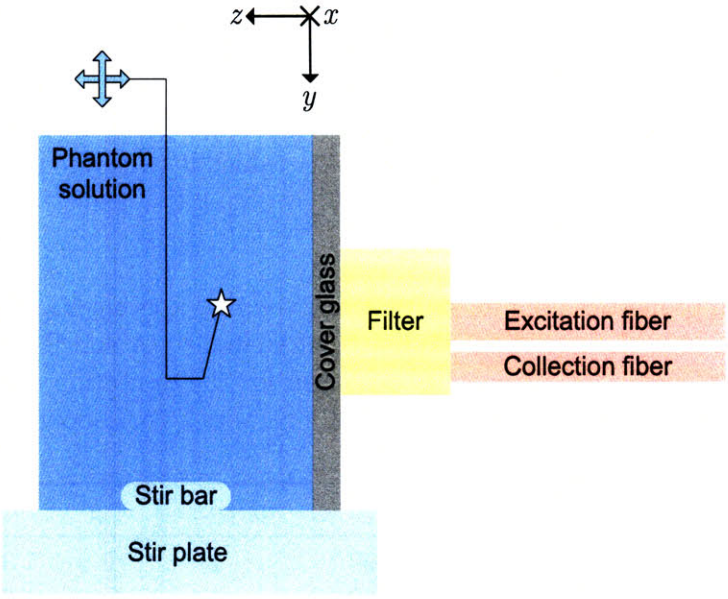


Figure 6-5: Sampling volume experiment schematic. A small diamond chip mounted on a needle was immersed in a scattering solution and translated through the sampling volume of a two fiber Raman probe.

The capped tube was turned upside, and the top was cut off to form a flat-bottom cup. A notch was cut into the side of the cup and replaced with a cover glass (Fisherbrand Premium, Thermo Fisher Scientific), which was bonded with epoxy and then sealed with silicone to prevent leakage. The wall thickness of the cover glass was measured to be 150 μm.

To keep the phantom solution suspended during the experiments, a stir bar was placed in the bottom of the container, and the container was mounted on a magnetic stir plate (MS-7, Tri-R Instrument, Inc.; Rockville Centre, New York). The stir plate was then mounted on a tilt/rotation stage (485, Newport; Irvine, CA), which was used to align the container with the diamond scan axes.

The optical properties of the phantom solution were chosen to mimic human arterial tissue. A concise meta-analysis of the published optical property measurements for aorta[2] shows that optical property data is sparse in our wavelength region (800 – 1000 nm). Extrapolating from the published values, $\mu_a \sim 1$ cm^{-1}, $\mu_s \sim 100 - 300$ cm^{-1}, and $g \sim 0.9$. Based on these values, we created two phantom solutions, with the optical properties shown

[2] "Optical Properties of Aorta," by Steven Jacques, Oregon Medical Laser Center, http://omlc.ogi.edu/spectra/aorta/index.html

in Table 6.1. Optical properties are given at two wavelengths: 830 nm, the excitation wavelength, and 930 nm, which corresponds to the diamond Raman shift at 1332 cm^{-1}. The only difference between the two sets of optical properties, which we term "property set A" and "property set B," is that A has $\mu_s = 200$ cm^{-1}, which is twice the value of B ($\mu_s = 100$ cm^{-1}).

Table 6.1: Optical properties used in the sampling volume experiments. Two sets of optical properties were used in the sampling volume experiments; set A and B differ by a factor of 2 in the scattering coefficient. Properties are given for the excitation wavelength, 830 nm and the diamond Raman signal wavelength, 930 nm, which corresponds to a Raman shift of 1332 cm^{-1}.

Set	wavelength [nm]	μ_s [cm^{-1}]	μ_a [cm^{-1}]	g	n
A	830	200	1	0.9	1.33
A	930	160	1	0.88	1.33
B	830	100	1	0.9	1.33
B	930	80	1	0.88	1.33

Probe geometry

For the purpose of comparing simulated and experimental sampling volume, we used a simplified Raman probe design to reduce potential discrepancies between the experimental and simulation geometries. The probe consisted of only two optical fibers, a single-mode fiber for excitation and a multimode fiber for collection. While the Raman probes of Chapters 4 and 5 used identical multimode fibers for both excitation and collection, the simulation requires precise knowledge of the excitation beam geometry. Modeling multimode fiber propagation is nontrivial, especially given the short length of fiber (~2 m) used in a typical Raman probe and the poorly characterized beam properties of the diode laser source. Thus, to enable comparison of the experimental and simulation results, we used a single-mode excitation fiber (SM800-5.6-125, Fibercore Ltd; Southampton, UK), which has a well-defined Gaussian beam output, with a mode field diameter of 5.6 μm. The collection fiber was identical to the fiber used in the previously decribed Raman probes (100 μm core diameter, 0.26 NA).

Both fibers were bonded to the filter using UV curing epoxy (Norland 61, Norland Products; Cranbury, New Jersey). The distance between the two fibers was approximately 250 μm, and the collection fiber face was offset from the filter surface by ~50 μm.

The excitation source was a Ti:sapphire laser (Tsunami, Spectra-Physics; Mountain View, CA) tuned to 830 nm in cw mode. The laser output was bandpass filtered at 830 nm and then coupled into the single-mode excitation fiber. The collection fiber was directly coupled to the spectrometer. Fianlly, the probe was mounted on a 3-axis kinematic mount, which in turn was mounted on a 3-axis translation stage, giving enough degrees of freedom to align the probe surface with the cover glass.

Diamond

The diamond chip used in these experiments was specifically chosen for its regular shape (Figure 6-6). The chip was approximately a parellelpiped, with dimensions of approximately $\sim 90 \times 90 \times 40$ μm.

The pointed tip of a 26 gauge hypodermic needle was gently sanded to provide a flat surface onto which the small diamond chip was glued. The diamond chip was attached with UV curing epoxy, and cured with a UV lamp for approximately 1 hour. After curing, the needle tip was cut off from the luer part of the hypodermic needle and glued onto copper wire with a two-part, 5 minute epoxy (Loctite). The copper wire was affixed to a 3-axis translation stage to scan the diamond through the probe's 3D sampling volume. The diamond was oriented such the short dimension was in the z-direction.

6.3.2 Experimental procedure

Alignment

Due to the small size scale of the this experiment, proper alignment of the probe, the cover glass, and the diamond chip is crucial. Comparison of the experimentally measured and simulated sampling volumes requires accurate knowledge of the experimental coordinates, especially the z-coordinate. Thus the system must be aligned to ensure that as the diamond is scanned through the scattering solution, it traces out planes that are parallel to the cover glass. Furthermore, the probe must be positioned in contact with the cover glass.

A surgical microscope was positioned vertically above the container to aid in the alignment. To align the container's cover glass with the stage's x-y plane, the diamond tip was translated towards the cover glass until it was visually determined to be just in contact; this contact determination was repeated several times and was repeatable to within 30-40 μm.

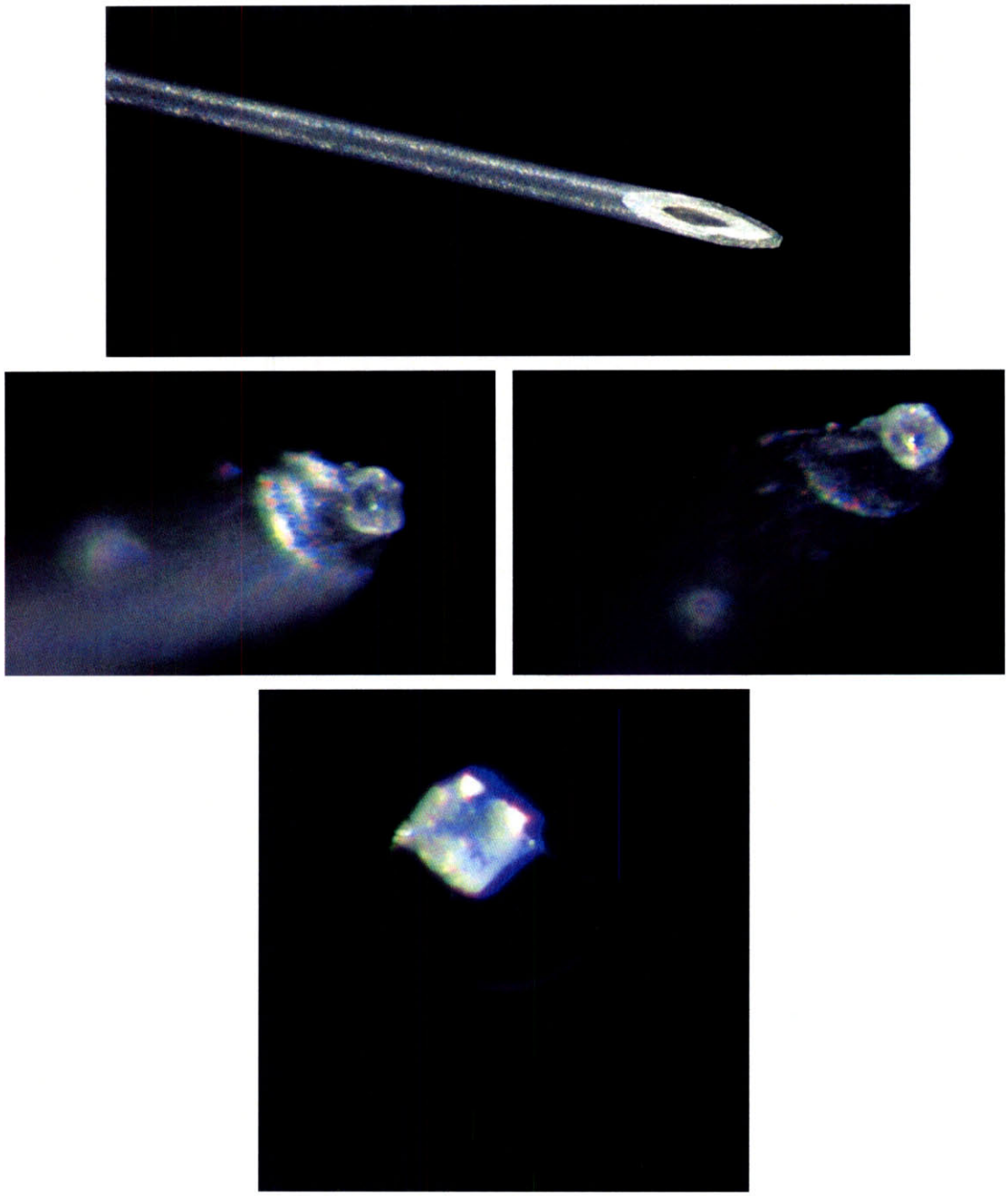


Figure 6-6: A diamond chip mounted on a hypodermic needle used for the sampling volume experiments. The diamond's dimensions are approximately $\sim 90 \times 90 \times 40$ µm.

The process was repeated at different locations on the x-y plane, and the tilt/rotation stage was adjusted until, for the corners of a 2×2 mm square, all four contact points were within 20 μm of each other. The probe was placed in contact with the cover glass by visualizing it with the surgical microscope and a magnifying glass.

Data acquisition

Prior to data acquisition, calibration spectra were acquired, including an acetaminophen spectrum for wavenumber calibration and a dark spectrum.

The diamond was positioned in contact with the cover glass, and then translated 100 μm away from the cover glass, in the positive z-direction, to avoid accidental loss of the diamond or scratches on the cover glass. The diamond was then manually translated in the x-y plane to maximize the Raman signal from the diamond; this position was taken to be $(x = 0, y = 0)$.

Raman spectra were acquired as the diamond was traced through a $200 \times 200 \times 200$ μm volume, with a step size of 20 μm in all dimensions. At each location, 10 frames of data were taken at 0.25 seconds per frame. Sampling volume measurements were repeated for both scattering phantom solutions at least two times.

Data processing

The Raman spectra were processed using the same preprocessing steps described in Chapter 2, including cosmic ray filtering, frame averaging, and PRNU correction. Because the probe illuminates the CCD differently from the benchtop system, a new PRNU correction factor was estimated, as described in Appendix A (Figure A-6).

A typical spectrum is shown in Figure 6-7. As expected, this spectrum exhibits the strong diamond band at 1332 cm^{-1}, as well as a broad background resulting from both the cover glass and fluorescence from the ink in the phantom solution. To extract the diamond band, a straight line was subtracted from the data between 1293 and 1366 cm^{-1}. The maximum value of the band was taken as the magnitude of the diamond signal.

6.3.3 Simulation procedure

All computations were performed on a computer running Windows XP, equipped with an Intel Core 2 Quad CPU (Q9300) and 2 GB of RAM.

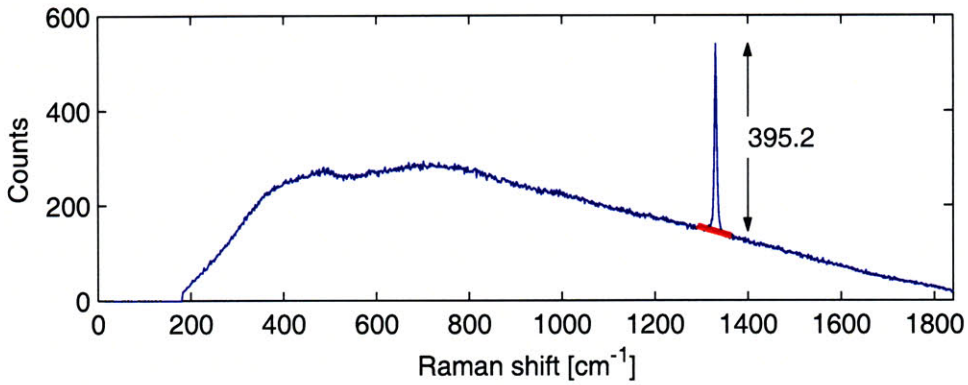


Figure 6-7: Data processing for diamond sampling volume experiments. The preprocessed spectrum is shown in blue. A linear background (red) was subtracted from the spectrum to determine the magnitude of the diamond Raman signal.

6.3.4 Zemax model

The experimental probe geometry was created in Zemax (Figure 6-8). The fibers were represented by concentric cylinders representing the core, cladding and buffer of the fiber, and the filter was a single piece of fused silica. In the fabricated probe, the collection fiber was offset from the filter by 50 μm, so this was added to the Zemax model. Additionally, epoxy was included in the model, surrounding the distal tip of the fibers.

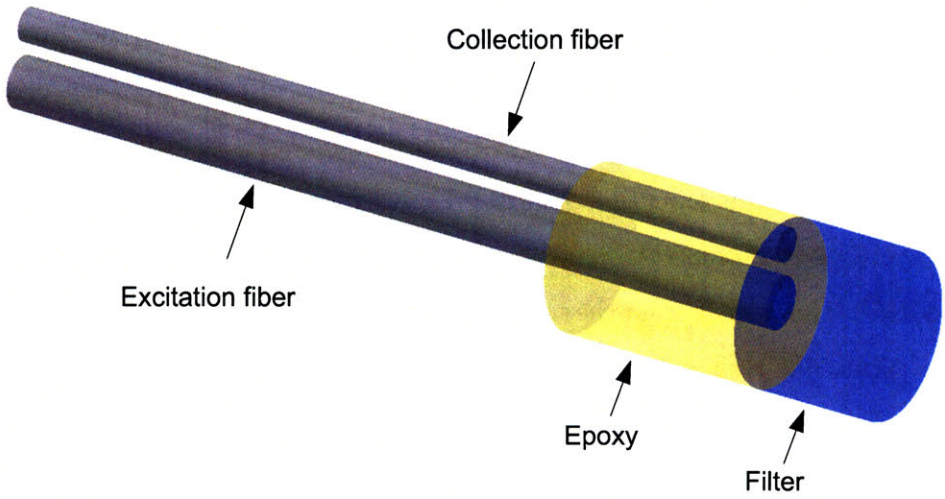


Figure 6-8: Zemax probe geometry for sampling volume experiments.

MCML model

MCML is programmed to allow the user to define a series of semi-infinite layers to represent the sample. To model the experimental sample, we defined a two-layer model, where the top layer is the cover glass and the bottom layer is the phantom solution. In the simulation, the thickness of the phantom layer was set to be arbitrarily large at 5 cm. The MCML model parameters are shown in Table 6.2.

Table 6.2: MCML model input parameters for the sampling volume simulations.

n	μ_a [cm^{-1}]	μ_s [cm^{-1}]	g	d [cm]	Comment
1.0					n for medium above (air)
1.515	0	0	-	0.015	Cover glass
1.33	μ_a	μ_s	g	5	Scattering medium (see Table 6.1)
1.0					n for medium below (air)

For each set of optical properties, the simulations were performed twice, using two scattering phase functions: the Mie phase function, which is accurate for the polystyrene microsphere solution, and the Henyey-Greenstein phase function, a commonly used approximation (Eq. (6.8)). At each scattering step of the simulation, a random deflection angle is generated from the chosen phase function, using the inverse cumulative density function (CDF) technique [123]. The CDF of the Henyey-Greenstein phase function is analytically invertible, but the CDF of the Mie phase function is not. Thus Monte Carlo simulations utilizing the Mie phase function require a table lookup step, increasing the total computation time.

For the 1 μm diameter microspheres used in the these validation experiments, the two phase functions are similar for both the excitation and collection wavelengths (Figure 6-9). The primary difference is that the Henyey-Greenstein phase function has a higher probability of forward scattering for very small deflection angles ($\phi < 12°$), while the Mie phase function has higher probability for $\phi \in (12°, 40°)$.

Step 1: Propagate excitation light to interface plane

While the simulation procedure in Section 6.2.1 calls for the use of Zemax to propagate from the excitation laser to the probe/sample interface plane, the experimental probe geometry was simple enough that Zemax was not necessary. The output of the single mode fiber is a

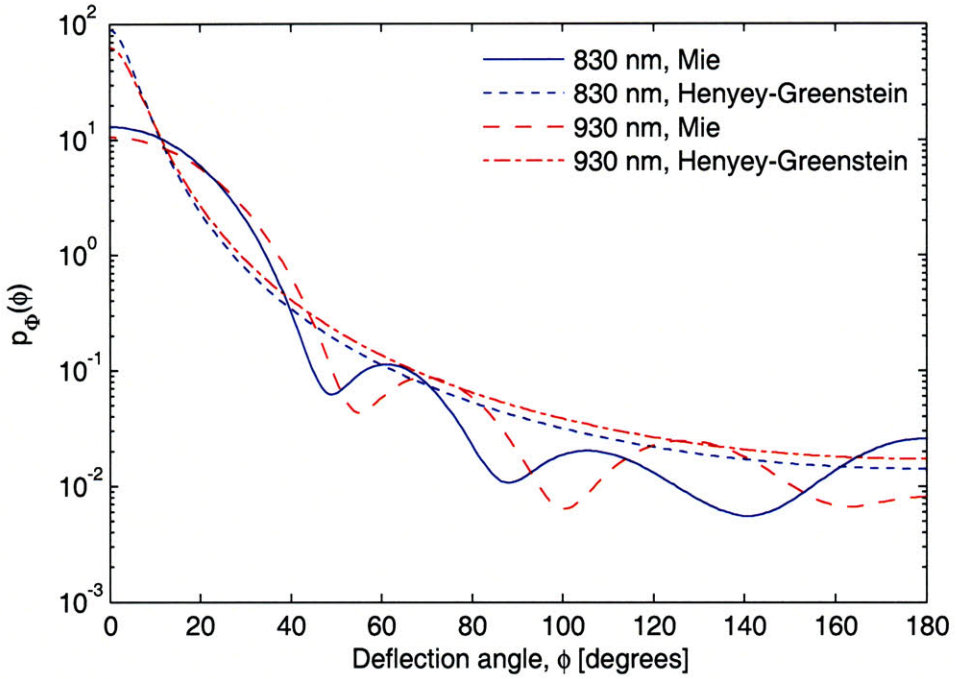


Figure 6-9: Mie and Henyey-Greenstein phase functions for 1 μm diameter polystyrene microspheres in water.

Gaussian beam with mode field diameter 5.6 μm and numerical aperture 0.12, which then travels through the optical filter to the interface plane. Using ABCD matrix analysis or divergence analysis for Gaussian beams [124], we determine that at the interface between the probe and the coverglass, the beam waist is $w = 39$ μm.

Step 2: Propagate excitation light into tissue

MCML was used to simulate the propagation of the excitation light within the tissue. Two million input rays were generated, each with unit weight, to simulate a Gaussian beam with the calculated beam waist. Each ray's position was generated randomly from a 2D Gaussian distribution:

$$I(\rho) = \frac{2P}{\pi w^2} \exp\left[-\frac{2\rho^2}{w^2}\right], \tag{6.18}$$

where $\rho = \sqrt{x^2 + y^2}$ and P is the beam power, set to unity. The initial propagation direction of each ray was determined based on the ray's radial position, ρ. The initial

deflection angle, ϕ, was calculated as

$$\sin(\phi) = \frac{\rho}{R}, \tag{6.19}$$

and the azimuthal angle, θ was defined to be radially outward,

$$\tan(\theta) = \frac{y}{x}. \tag{6.20}$$

The modified MCML was used to propagate the excitation rays, using the parameters in Table 6.2. Absorption in each $10\times10\times10$ μm voxel was recorded, yielding $E_{\text{Raman,gen}}(x, y, z)$. The computation time for step 2 was approximately 30 minutes using the Henyey-Greenstein phase function, and 45 minutes using the Mie phase function.

Step 3: Propagate isotropically emitted Raman photons to interface plane

We defined Raman point sources as being $10 \times 10 \times 10$ μm voxels. All point sources were centered at $x = 0$ and $y = 0$. Because the cover glass occupied the first 150 μm of the sample geometry, point source voxels were generated to span $z = 150$ μm to $z = 560$ μm, for a total of 41 point sources. For each point source, two million photons were initialized, each with a uniformly distributed random position within the voxel and an isotropically distributed propagation angle. The photons were propagated to the interface plane using the modified MCML. Because MCML uses only one CPU core, the 41 point sources were split into four groups, and each group was then assigned to one core. The total computation time for step 3 was approximately $3 - 4$ hours using the Henyey-Greenstein phase function and $7 - 9$ hours using the Mie phase function.

Step 4: Propagate Raman photons from interface plane through probe

Because this probe geometry consists only of a flat filter, the collection profile for the collection fiber, $E_{\text{Raman,coll}}(x, y, z)$, is radially symmetric with respect to the center of the collection fiber. We exploited this radial symmetry to reduce the computation, by only propagating Raman photons from isotropic point sources located in a single quadrant under the collection fiber. Additionally, because $E_{\text{Raman,coll}}(x, y, z)$ is a smoothly varying function, we further reduced the computation by shifting the point sources in x and y by 20 μm steps, and then interpolating the results to determine $E_{\text{Raman,coll}}(x, y, z)$ on the desired $10\times10\times10$

μm grid. The total computation time for step 4 was approximately 12 hours for each phase function.

Step 5: Calculate the sampling volume

Finally, MATLAB was used to multiply the excitation profile from step 2 by the collection profile from step 4, to determine the sampling volume profile, $E_{\text{Raman,sampling volume}}(x, y, z)$. To compare the simulation results with the the experimental results, the finite size of the diamond was accounted for by convolving the calculated sampling volume with a matrix of ones corresponding to the size of the diamond. Computation time for step 5 was approximately 15 minutes for each phase function.

6.4 Results

6.4.1 Simulation results

The simulation results are shown in Figures 6-10 (property set A, $\mu_s = 200$ cm^{-1}, Henyey-Greenstein phase function), 6-11 (property set B, $\mu_s = 100$ cm^{-1}, Henyey-Greenstein phase function), 6-12 (property set A, Mie phase function), and 6-13 (property set B, Mie phase function), with identical colorscales to aid in the comparison of the results. The Raman source, $E_{\text{Raman,gen}}(x, y, z)$, is symmetric about the excitation fiber's central axis, so cross-sections of the profile are shown in (Figures 6-10(a), 6-11(a), 6-12(a), and 6-13(a)). The numerical values shown are not physically relevant. For both phase functions, within the top 50 μm of the phantom, the Raman source profile for property set A is approximately 1% greater than for property set B. Furthermore, the source profile for property set B extends to a deeper depth (50% of the maximum value around 280 μm) than for property set A (50% around 250 μm). These results are explained by the difference in the scattering coefficient between property sets A and B. Because property set A has a higher value for μ_s, photons travel a shorter distance before being absorbed, which increases the number of absorbed photons near the top of the phantom. The lower value of μ_s in property set B allows the photons to propagate further without changing direction, which explains B's narrower, deeper Raman source profile. Although the Mie and Henyey-Greenstein phase functions give qualitatively similar results, at deeper depths, the Raman source profile calculated using the Henyey-Greenstein function is up to 40% higher than the Mie function profile.

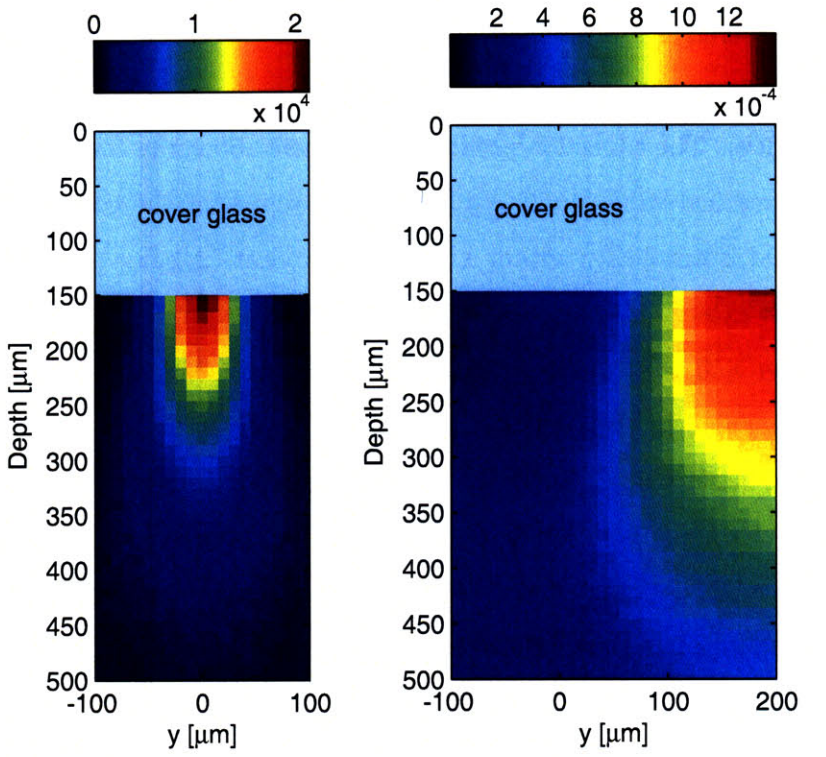
The discrepancy increases with increasing depth.

The Raman collection profiles, $E_{\text{Raman,coll}}(x, y, z)$, are shown in Figures 6-10(b), 6-11(b), 6-12(b), and 6-13(b). The collection profile is symmetric about the collection fiber's central axis, and the numerical values in $E_{\text{Raman,coll}}(x, y, z)$ represent the collection efficiency for photons generated isotropically within a voxel centered at (x, y, z). The collection profile for property set B extends deeper into the phantom than for property set A, and the overall shape is narrower and less rounded than A. Additionally, the profile for property set A is higher in magnitude in the central region of the fiber, at all depths. Again, the Mie and Henyey-Greenstein phase functions yield qualitatively similar results. Quantitatively, within the top ~100 μm, the Mie phase function has a lower magnitude collection profile, but at deeper depths, the Mie function yields up to 10% higher collected signal magnitude.

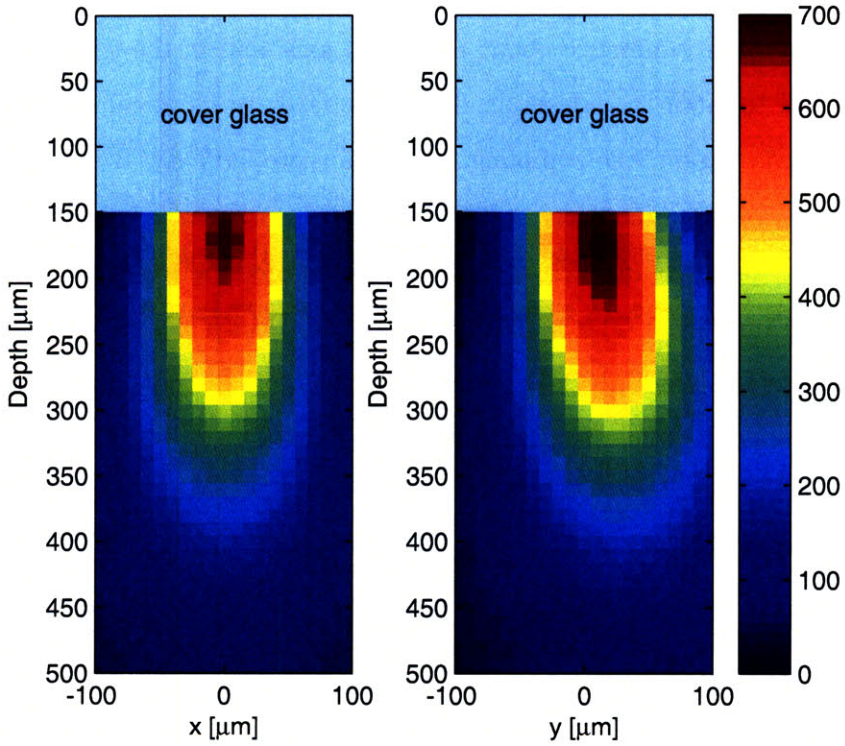
Finally, the Raman sampling volumes, $E_{\text{Raman,sampling volume}}(x, y, z)$, are shown in Figures 6-10(c), 6-11(c), 6-12(c), and 6-13(c), after convolution to account for the diamond size and shape. The numerical values are not physically meaningful. The sampling volume is not symmetric, so slices representing the x-plane and y-plane are shown. In the x-plane, the sampling volume is symmetric about the depth axis, and in the y-plane, the sampling volume is skewed towards the collection fiber; this leads to elliptical contours in each z-plane (not shown). Again, the sampling volume for property set B extends deeper into the tissue and is narrower than for property set A, which is shallower and more rounded. The convolved sampling volumes show qualitative differences between the Mie and Henyey-Greenstein phase functions. The Mie sampling volume decays faster and has a maximum at a deeper depth than the Henyey-Greenstin sampling volume, possibly due to the smaller forward scattering of the Mie phase function.

6.4.2 Experimental results

The experimental sampling volume results are shown in Figures 6-14 and 6-15, with identical colorscales to aid in comparison. The numerical values correspond to CCD counts of the magnitude of the diamond spectrum. At each depth, the contours of the diamond signal exhibit an elliptical shape, as predicted by the simulations. Serendipitously, the major and minor axes of the ellipses also lie along the experimental system's x and y axes, so no rotation of the experimental dataset was necessary to compare the simulated results. For both optical property sets, experimental data was acquired within the same volume, a

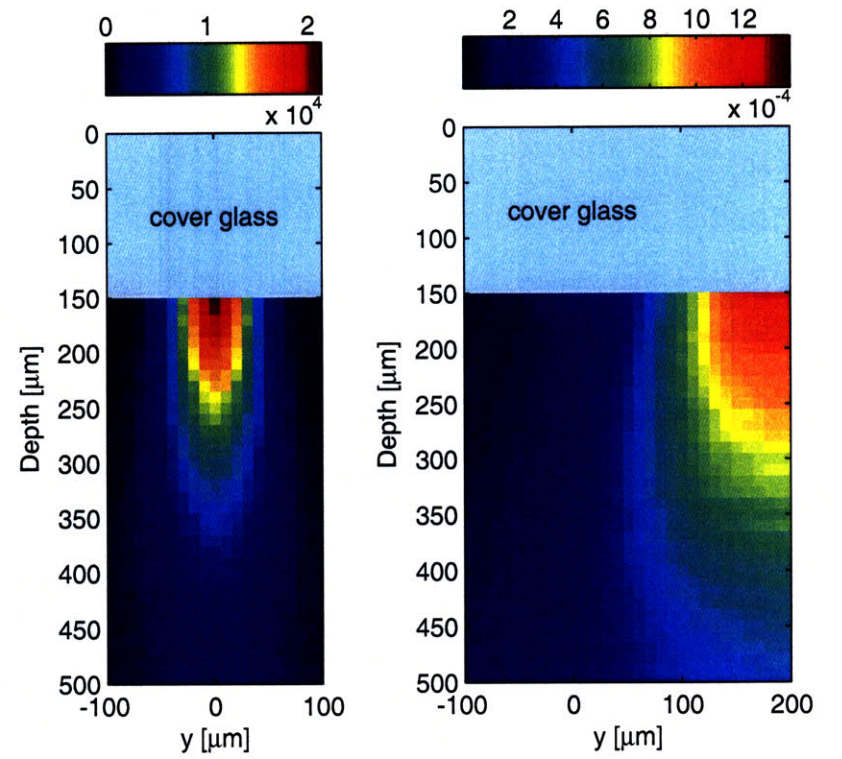


(a) $E_{\text{Raman,gen}}(x, y, z)$ (b) $E_{\text{Raman,coll}}(x, y, z)$

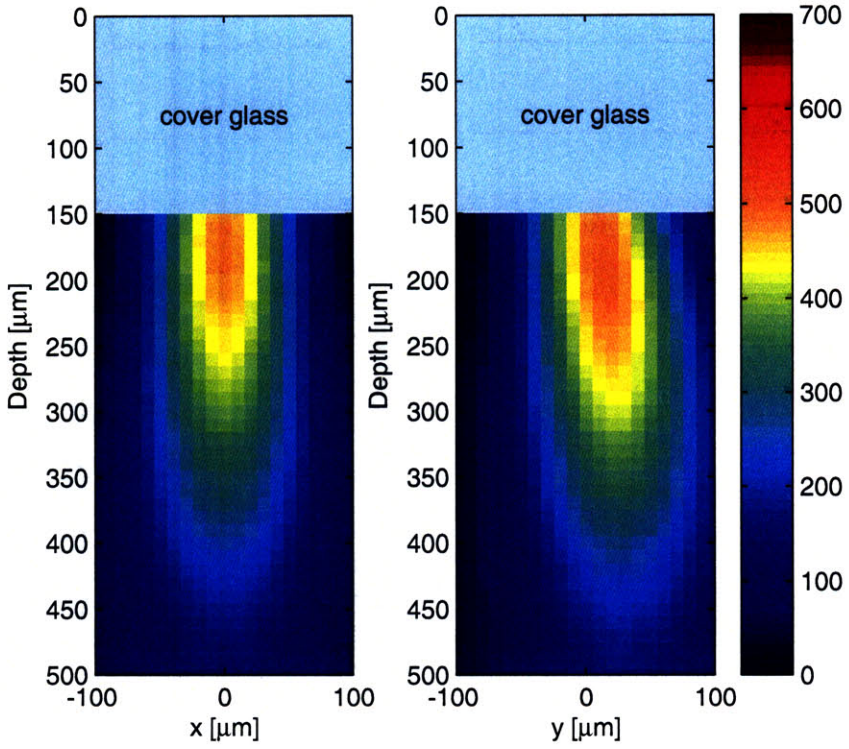


(c) $E_{\text{Raman,sampling volume}}(x, y, z)$

Figure 6-10: Simulated sampling volume for property set A ($\mu_s = 200$ cm^{-1}), with Henyey-Greenstein phase function. (a) Raman source profile, (b) Raman collection profile, and (c) Raman probe sampling volume, after convolution.

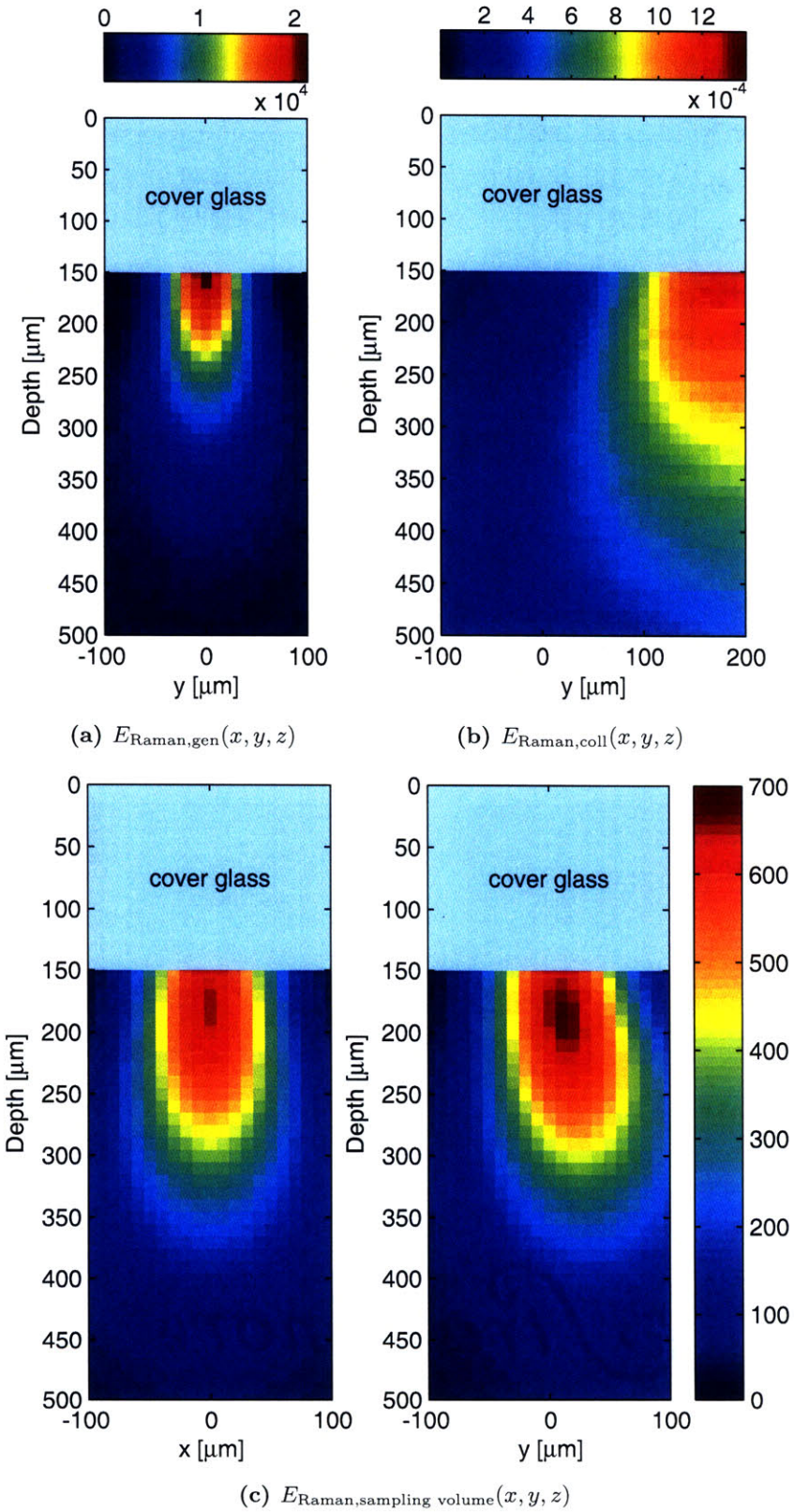


(a) $E_{\text{Raman,gen}}(x, y, z)$ (b) $E_{\text{Raman,coll}}(x, y, z)$



(c) $E_{\text{Raman,sampling volume}}(x, y, z)$

Figure 6-11: Simulated sampling volume for property set B ($\mu_s = 100$ cm^{-1}), with Henyey-Greenstein phase function. (a) Raman source profile, (b) Raman collection profile, and (c) Raman probe sampling volume, after convolution.

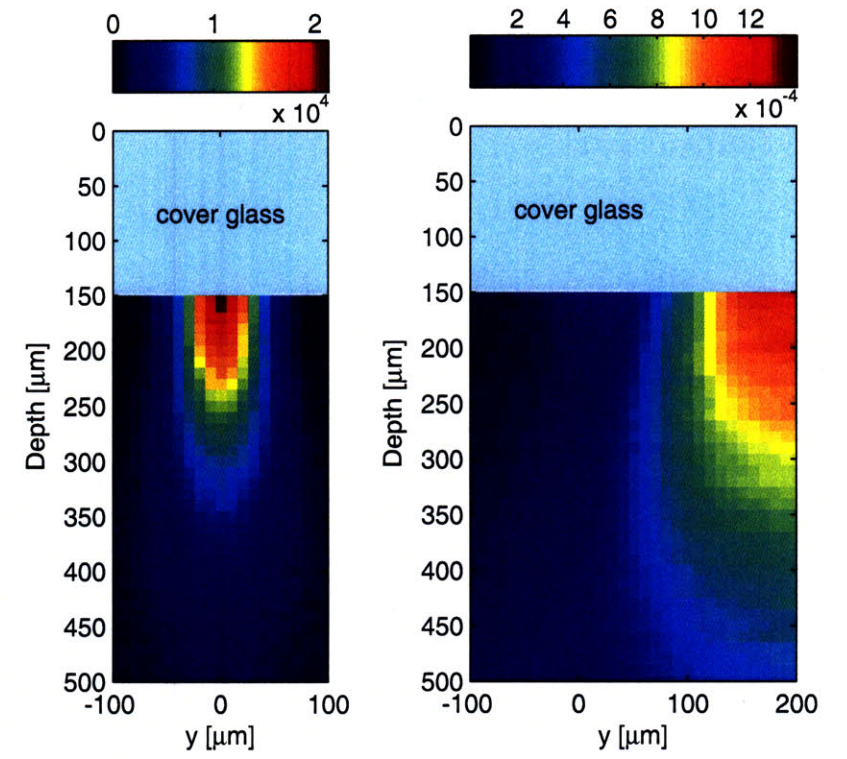


(a) $E_{\text{Raman,gen}}(x, y, z)$

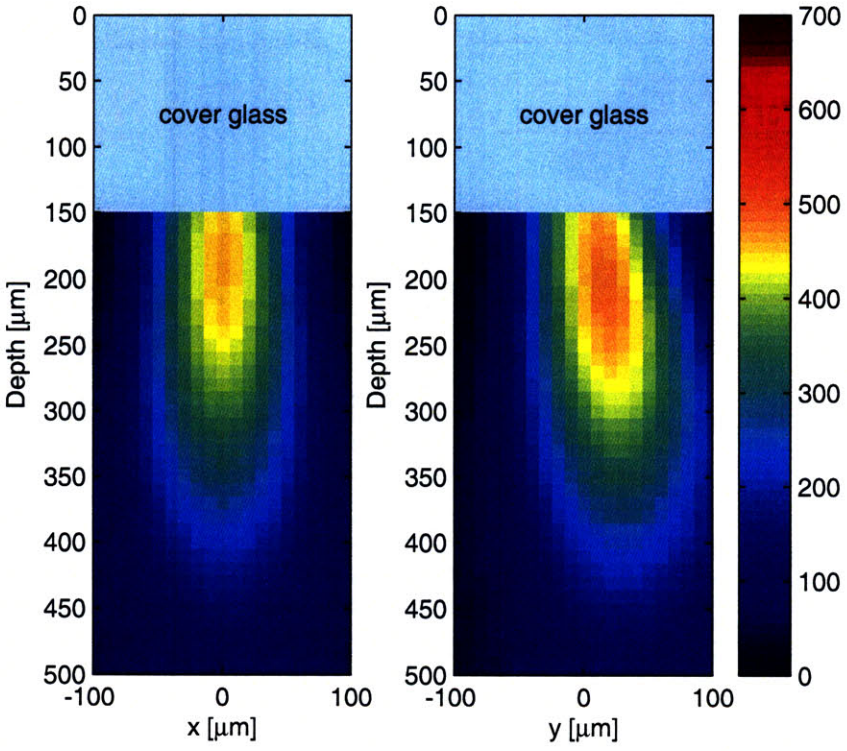
(b) $E_{\text{Raman,coll}}(x, y, z)$

(c) $E_{\text{Raman,sampling volume}}(x, y, z)$

Figure 6-12: Simulated sampling volume for property set A ($\mu_s = 200$ cm^{-1}), with Mie phase function. (a) Raman source profile, (b) Raman collection profile, and (c) Raman probe sampling volume, after convolution.

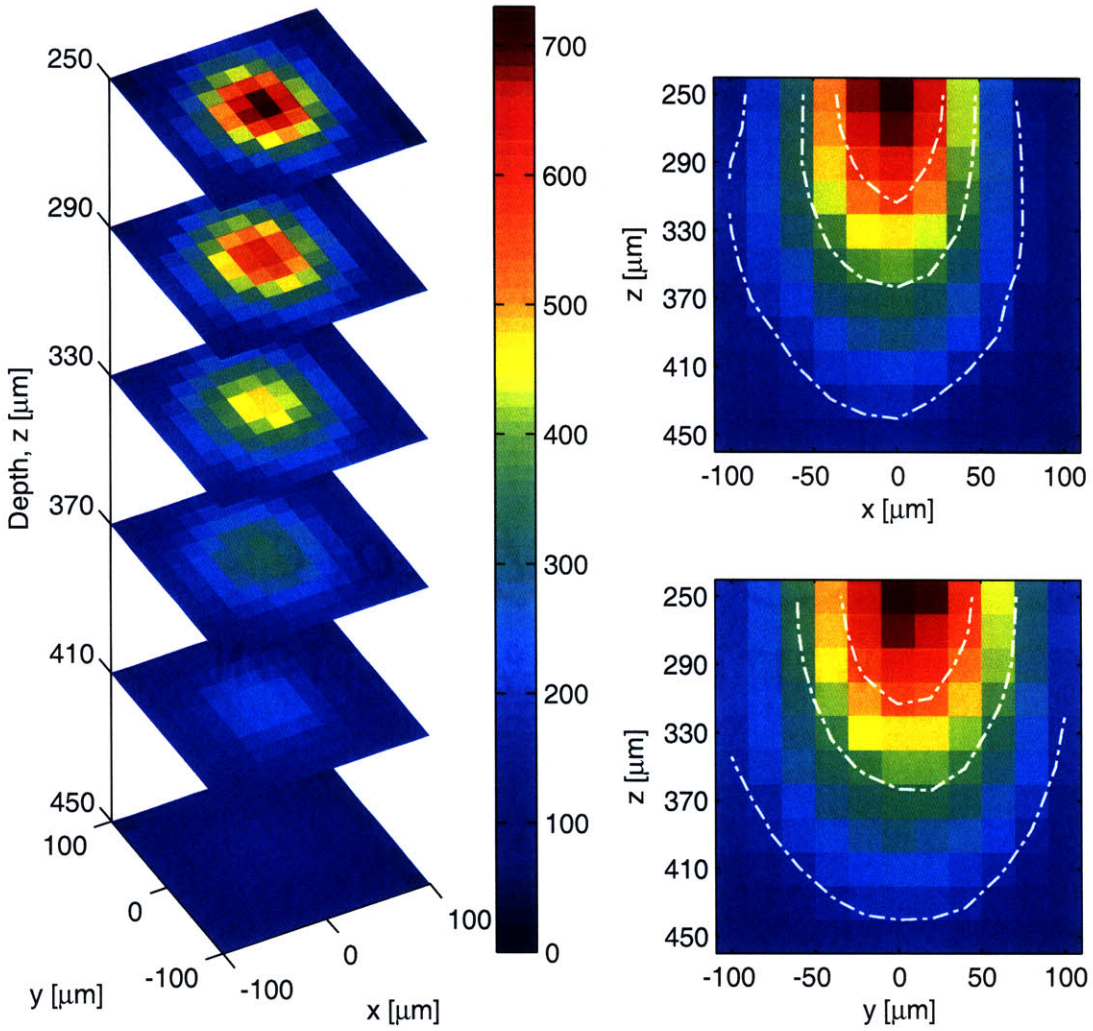


(a) $E_{\text{Raman,gen}}(x, y, z)$ (b) $E_{\text{Raman,coll}}(x, y, z)$



(c) $E_{\text{Raman,sampling volume}}(x, y, z)$

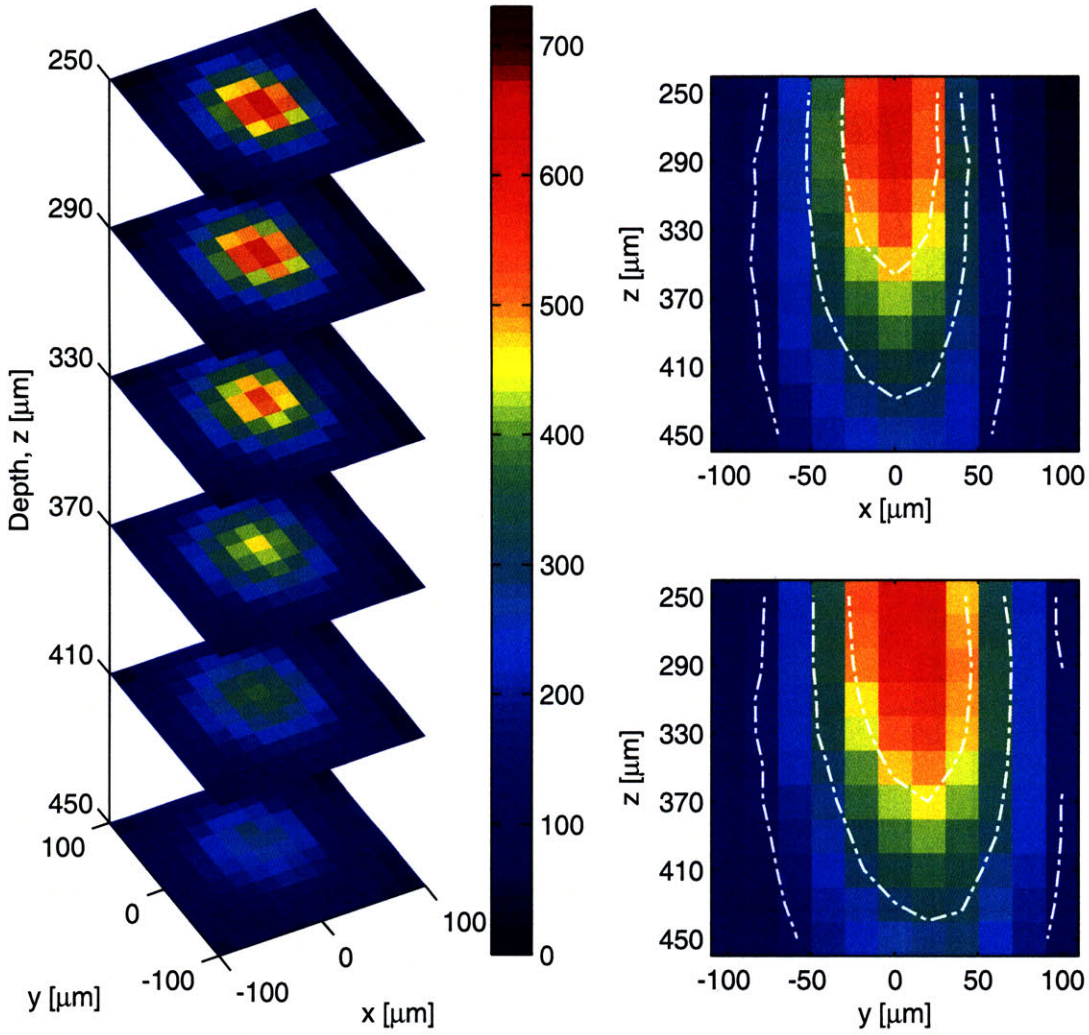
Figure 6-13: Simulated sampling volume for property set B ($\mu_s = 100$ cm^{-1}), with Mie phase function. (a) Raman source profile, (b) Raman collection profile, and (c) Raman probe sampling volume, after convolution.



(a) Multiple depths

(b) x- and y-planes

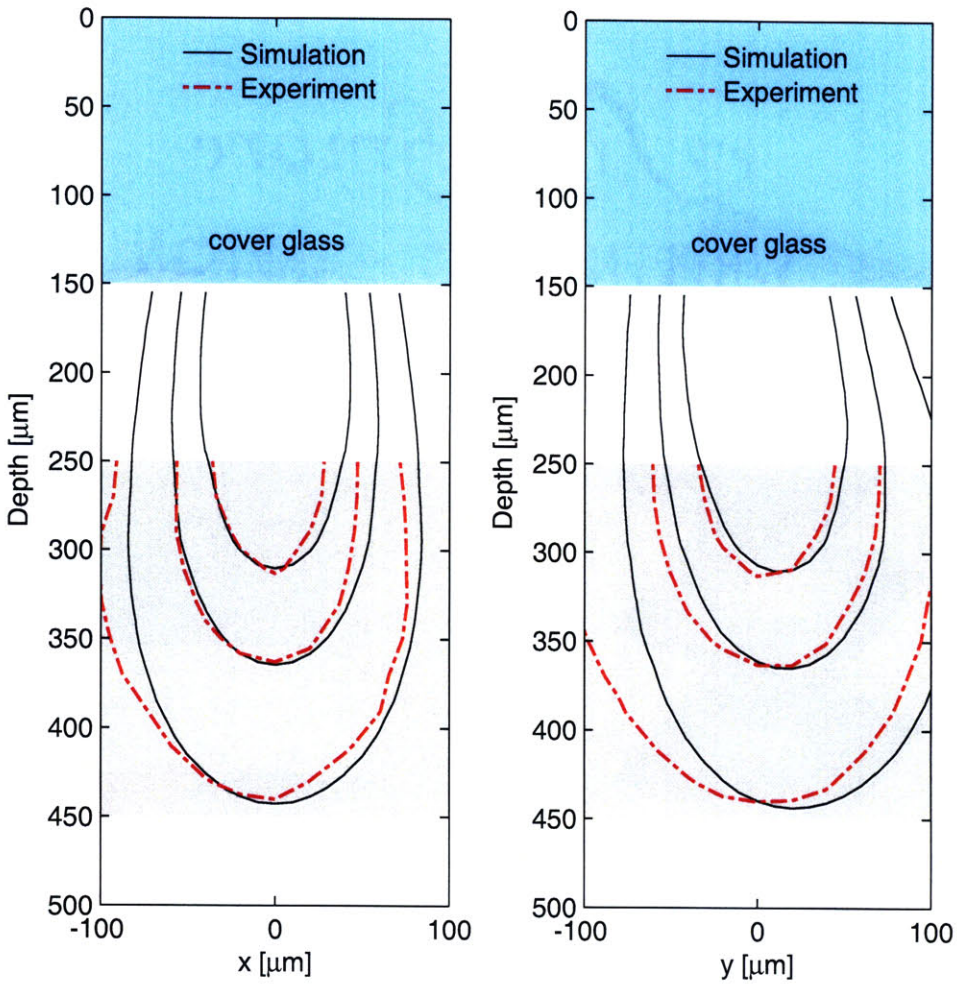
Figure 6-14: Experimentally measured sampling volume for property set A ($\mu_s = 200\ \mathrm{cm}^{-1}$). (a) Diamond signal (CCD counts) for multiple depths. (b) Diamond signal in the x-plane and y-plane. White isocontour lines represent 75%, 50%, and 25% of the maximum value.



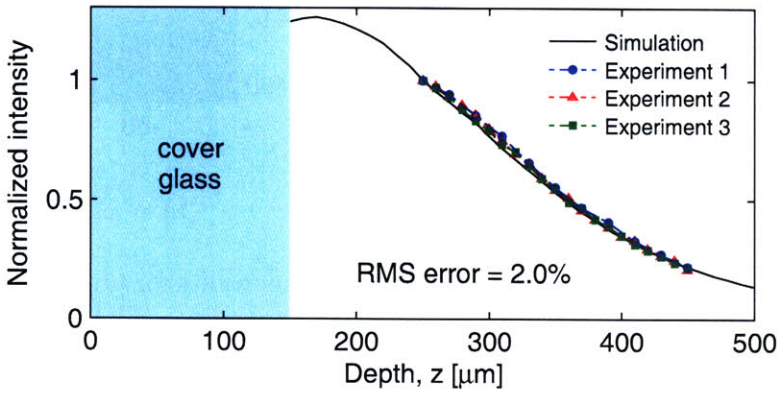
(a) Multiple depths

(b) x- and y-planes

Figure 6-15: Experimentally measured sampling volume results for property set B ($\mu_s = 100$ cm^{-1}. (a) Diamond signal (CCD counts) for multiple depths. (b) Diamond signal in the x-plane and y-plane. White isocontour lines represent 75%, 50%, and 25% of the maximum value.

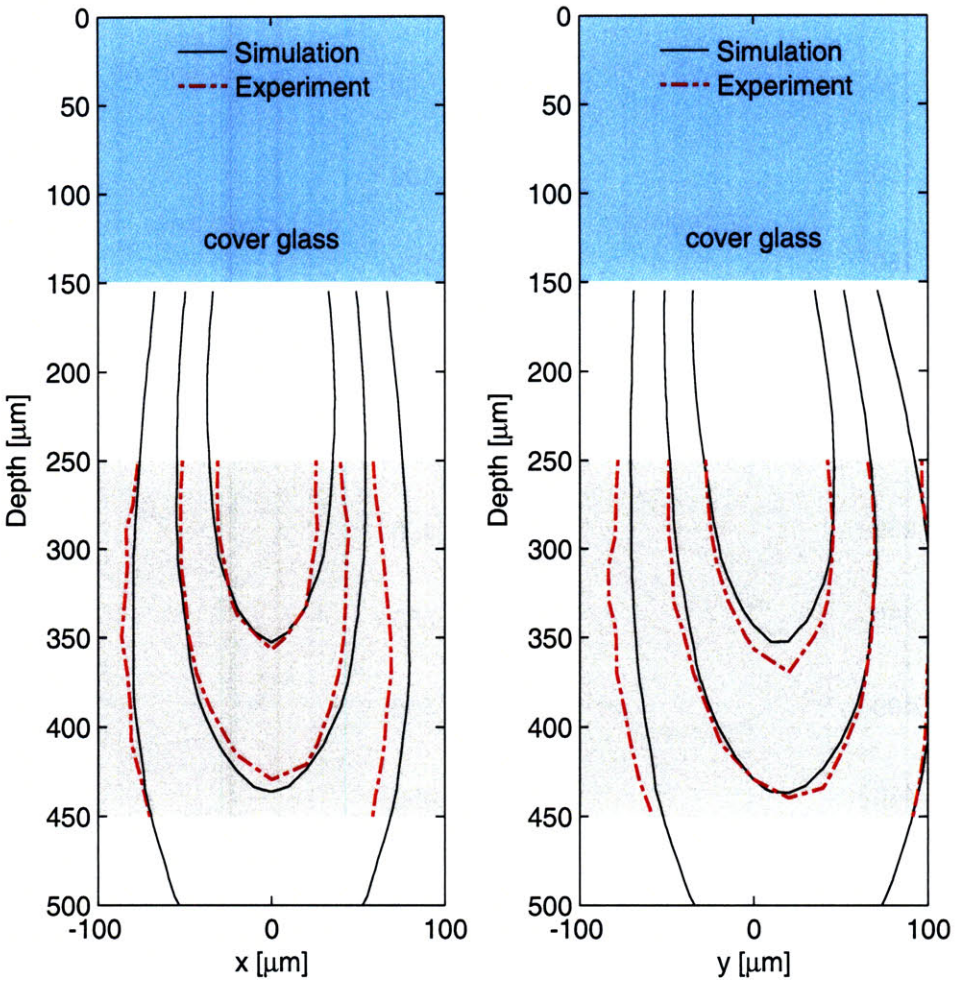


(a) Isocontours at 75%, 50%, and 25% of the maximum value at $z = 250$ micron.

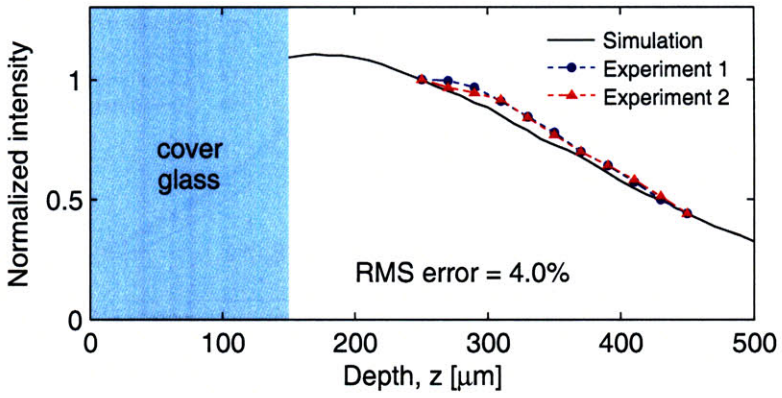


(b) $E_{\text{Raman,sampling volume}}(x = 0, y = 0, z)$

Figure 6-16: Comparison of experimental and simulated sampling volumes for property set A ($\mu_s = 200$ cm^{-1}), with Henyey-Greenstein phase function.

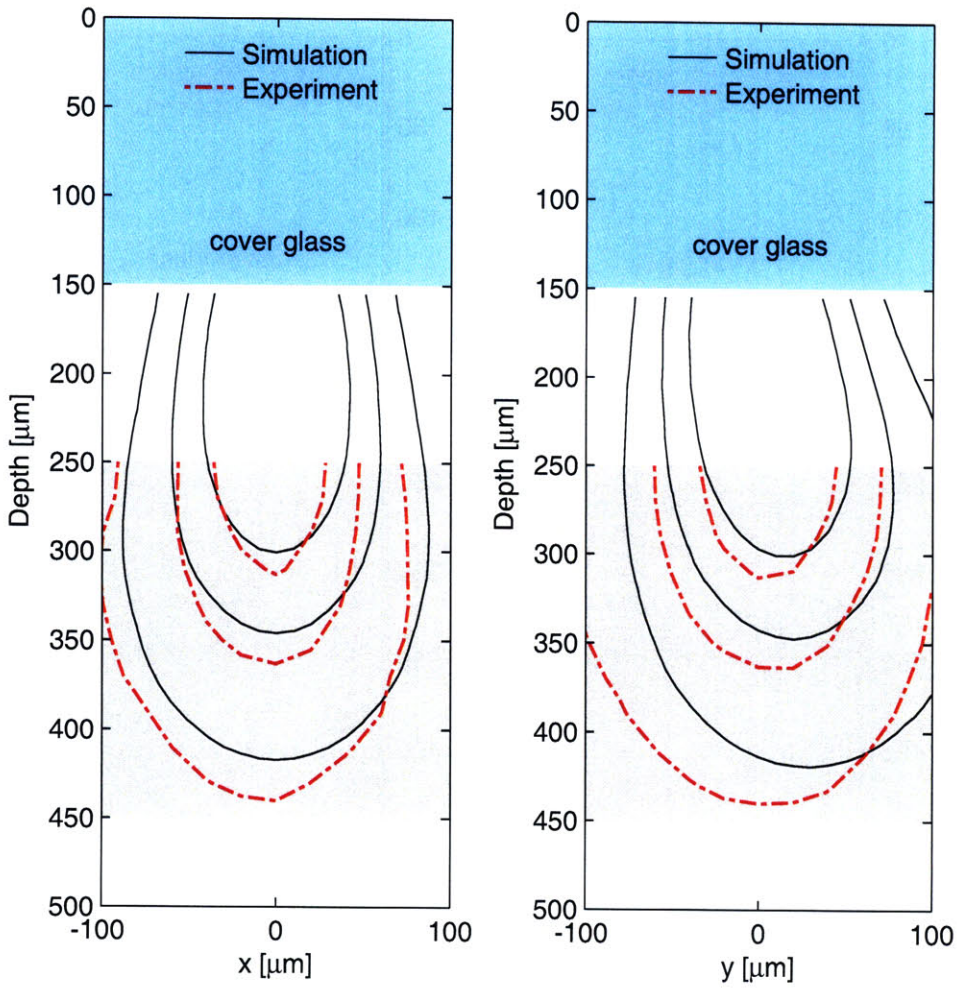


(a) Isocontours at 75%, 50%, and 25% of the maximum value at $z = 250$ micron.

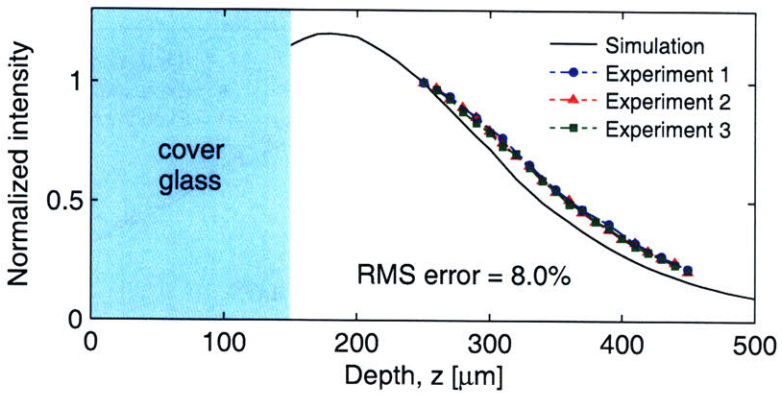


(b) $E_{\text{Raman,sampling volume}}(x = 0, y = 0, z)$

Figure 6-17: Comparison of experimental and simulated sampling volumes for property set B ($\mu_s = 100$ cm^{-1}), with Henyey-Greenstein phase function.

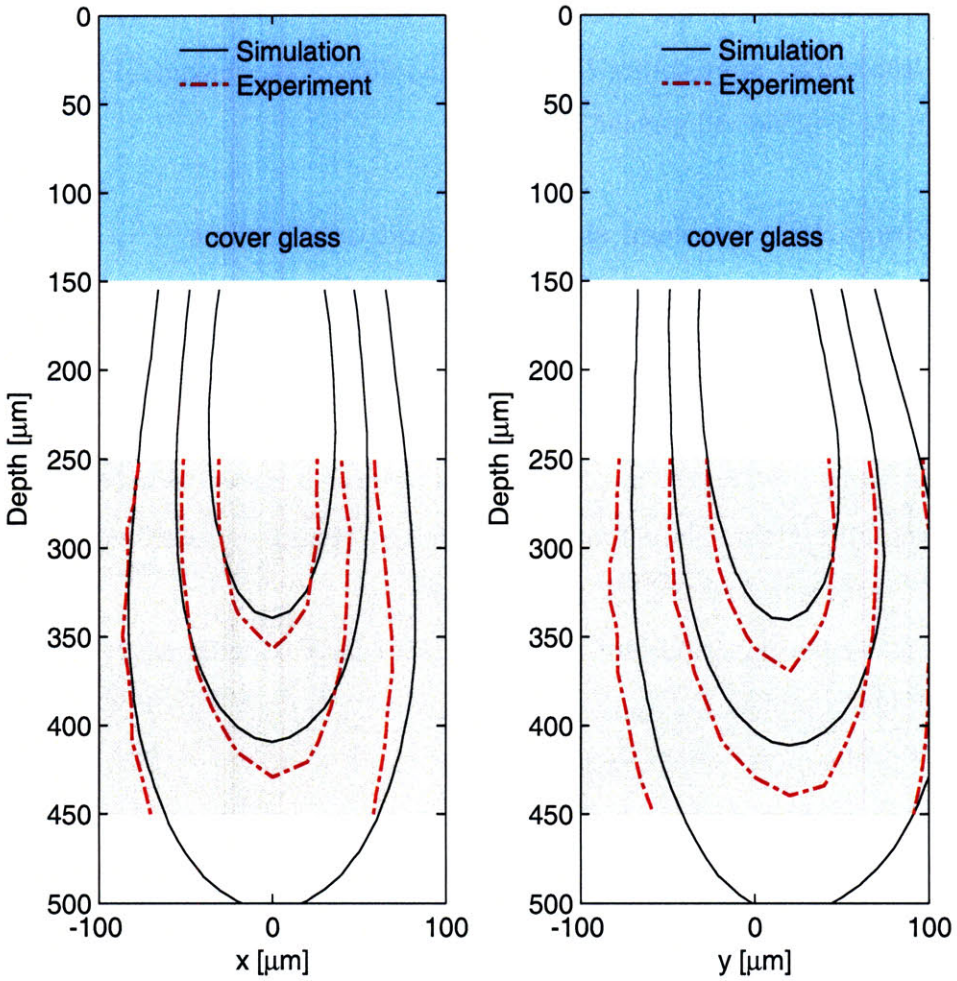


(a) Isocontours at 75%, 50%, and 25% of the maximum value at $z = 250$ micron.

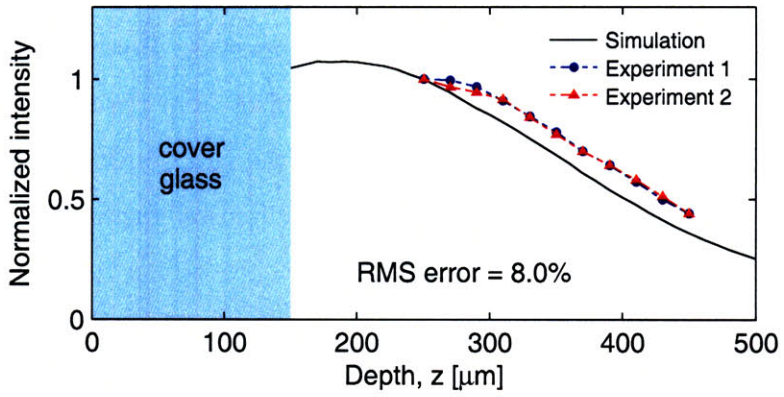


(b) $E_{\text{Raman,sampling volume}}(x = 0, y = 0, z)$

Figure 6-18: Comparison of experimental and simulated sampling volumes for property set A ($\mu_s = 200$ cm^{-1}), with Mie phase function.



(a) Isocontours at 75%, 50%, and 25% of the maximum value at $z = 250$ micron.



(b) $E_{\text{Raman,sampling volume}}(x = 0, y = 0, z)$

Figure 6-19: Comparison of experimental and simulated sampling volumes for property set B ($\mu_s = 100$ cm^{-1}), with Mie phase function.

$200 \times 200 \times 200$ μm cube, with a step size of 20 μm. At the top of the measured volume ($z = 250$ μm), the sampling volume in property set A is greater in magnitude than in B, as predicted by the simulations. Additionally, A's sampling volume decays faster in depth than B's. Isocontours representing 75%, 50%, and 25% of the maximum signal value are displayed in the x-plane and y-plane images.

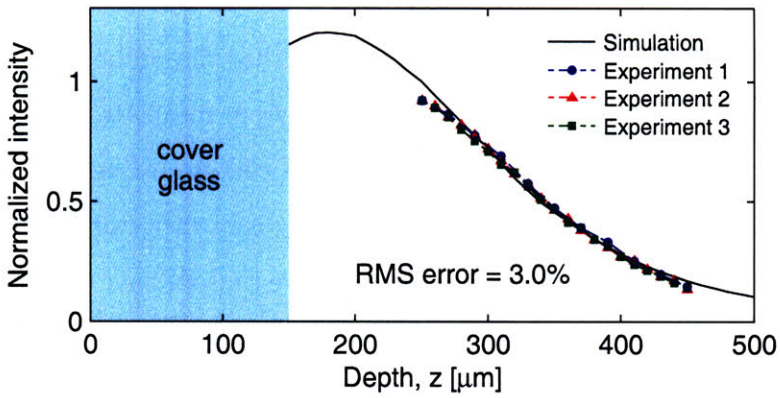
6.4.3 Comparison between simulation and experiment

The diamond signal isocontours can be used to qualitatively assess the agreement between the experimentally and simulation derived sampling volumes (Figures 6-16(a), 6-17(a), 6-18(a), and 6-19(a)). Contours were calculated as 75%, 50%, and 25% of the maximum signal at the depth $z = 250$ μm. In all cases, good agreement is seen between the simulation and experimental contour values, though the Henyey-Greenstein phase function appears to yield better agreement than the Mie phase function.

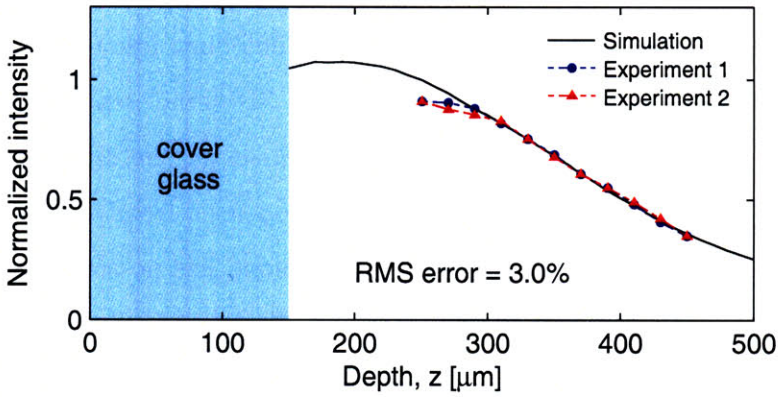
For quantitative comparison, we plot the simulated and experimentally measured one-dimensional depth profile, $E_{\text{Raman,sampling volume}}(x = 0, y = 0, z)$ (Figures 6-16(b), 6-17(b), 6-18(b) and 6-19(b)). All values were normalized to the value at $z = 250$ μm. The results from three experimental repetitions of property set A, taken within the same experimental session, and two repetitions of property set B, taken on different days, demonstrate that the experiments are highly repeatable. Simulations using the Henyey-Greenstein phase function yield better agreement with the experimental sampling volume, with an RMS error of 2% for property set A, and 4% for property set B. In comparison, the Mie phase function results have an RMS error of 8% for both sets of optical properties. However, for $z > 300$ μm, the slope of the experimental results appears to be offset from the simulation curve. By normalizing the results to the value at $z = 310$ μm, the RMS error is reduced to 3% for both property sets (Figure 6-20). This phenomenon is discussed further in the next section.

6.5 Discussion

In this chapter, we have developed a Monte Carlo model for Raman scattering, and used it to simulate the sampling volume of a Raman probe. The simulation technique considers the propagation of both the excitation light and the returning Raman light, which is important because both beam paths may be affected by the addition of distal optics on the probe tip.



(a) Property set A, with Mie phase function



(b) Property set B, with Mie phase function

Figure 6-20: Experimental and simulated (Mie phase function) sampling volume. The experimental curves have been offset, effectively normalizing to the value at $z = 310\mu m$, giving good agreement for $z > 300\mu m$.

Furthermore, the model treats the excitation and collection beam paths separately, so it is only necessary to simulate the isotropic emission of Raman photons once for a given set of optical properties, which reduces the total computation time.

We have experimentally validated this Raman optical simulation technique by using a diamond chip mounted on a needle to measure the sampling volume of a two-fiber Raman probe. The small diamond chip was swept through a scattering phantom solution, while the magnitude of the diamond Raman signal was measured. The phantoms were aqueous solutions of polystyrene beads and india ink, with two distinct sets of optical properties chosen within the range of optical properties for aorta. For both phantoms, the simulated and experimentally measured Raman sampling volumes show good agreement.

There are several potential explanations for the discrepancies between the simulated and experimental sampling volume results. The most important is geometry discrepancies.

The Raman probe is small — the outer diameter of the optical fibers in the probe is approximately 200 μm. At this size scale, it is likely that the probe was not fabricated to the desired specifications. Indeed, while both fibers were meant to be bonded in contact with the filter surface, in the fabricated probe, the collection fiber was offset from the filter surface by 50 μm. Another potential discrepancy is in the positioning of the diamond relative to the container's cover glass. While we believe that we have accurately determined the z location of the diamond, this measurement could be off by up to ~20 μm.

Another important difference between the simulations and experiments lies in the way the diamond was modeled in the simulation. We simulated the Raman sampling volume for a medium with homogeneous optical properties, and then accounted for the diamond by convolving the simulation results with a uniform 3D kernel representing the size and shape of the diamond. However, in the experimental setup, the diamond is a $90 \times 90 \times 40$ μm inclusion with distinctly different optical properties from the phantom solution; the index of refraction for diamond is ~2.4 and diamond has no scattering or absorption at this wavelength. Additionally, the diamond chip does not have regular facets, further complicating the optical geometry.

We have performed simulations using both the Henyey-Greenstein and Mie scattering phase functions. While the Mie phase function is accurate for the polystyrene bead solution, it is more computationally intensive, and thus the Henyey-Greenstein approximation is often used. Compared to the Henyey-Greenstein simulations, the sampling volume simulations using the Mie phase function better match the trends seen in the experimental sampling volumes for $z > 300$ μm, which is expected since the Mie phase function more accurately describes scattering in the phantom solutions. This effect is most noticeable in property set B; in property set A, simulation results from both phase functions match the experimental results well. This is likely due to the overall increased scattering in A, where the collected photons have undergone so many scattering events that the minor differences in the two phase functions become less important. However, the Mie phase function simulations exhibit a discrepancy with the experimental results at shallower depths, possibly because at shallower depths, the diamond inclusion boundary effect described above is dominant. Again this boundary effect is less prominent in property set A than in B, due to the increased average number of scattering events in A.

The Raman optical simulation technique may be a valuable tool for evaluating and

optimizing Raman catheter designs. There are several possible performance metrics for choosing the "optimal" design, the most obvious being some measure of collection efficiency or total Raman intensity. However, sampling volume is also important, and it is possible that targeting a specific Raman sampling volume may be an important design specification, in addition to maximum collection efficiency.

For example, in a scanning catheter, the catheter's sampling volume will determine the sampling rate needed to effectively screen a large segment of the artery. The sampling volume is likely to be important diagnostically as well; the interpretation of the Raman spectrum originating from a deep, narrow sampling volume is different from the interpretation of a shallower, wider sampling volume. The Raman optical simulation technique is an ideal platform for addressing these design questions.

There are several design parameters that can be optimized using the Raman simulation technique. For example, the optical fiber diameter and density directly affects the probe's sampling volume. Choosing fibers with increased numerical aperture may increase the amount of accepted Raman light. Furthermore, due to the Raman catheter's small size scale, the distal optical elements required to implement a side-viewing geometry are likely to be custom-fabricated, and the Raman simulation technique provides a means for designing such custom mirrors, lenses, and prisms.

Additionally, the Raman simulation technique can be used to determine fabrication tolerances for the catheter components. For example, because the Raman probe diameter is < 1 mm and the optical components are on a sub-millimeter size scale, misalignment during fabrication is likely. The simulation technique can be used to perturb the probe design through translations and tilts of the optical components. In this way, the simulations can be used to choose optical designs that are relatively simple to fabricate and robust to fabrication errors, giving high yield in the fabricated probes.

6.6 Conclusion

In this chapter, we have developed and experimentally validated a Raman optical simulation method that can be used to assess the sampling volume and collection efficiency of a Raman probe. The simulated and experimentally measured sampling volumes for two tissue-mimicking phantoms show good agreement. The simulation technique can be used to

design an optimal Raman catheter, based on the catheter's predicted sampling volume and collection efficiency over the range of expected catheter-tissue separation geometries and optical properties. Other potential uses of the simulation technique include designing custom optical elements for the catheter and determining fabrication tolerances in the catheter design.

Chapter 7

Conclusions and Future Directions

Atherosclerosis, a leading cause of death in industrialized nations, is an inflammatory disease of the arteries, characterized by the development of atheromatous plaques within the arterial wall. Rupture of these plaques and the resulting thrombosis can occlude flow and may ultimately result in myocardial infarction. The mechanism behind plaque progression and disruption is not well understood, but lesion structure, molecular and chemical composition, biomechanical behavior, and functional state are believed to be important. In this thesis, we have investigated Raman spectroscopy as a means for assessing the chemical composition of atherosclerotic plaques. Raman spectroscopy is an optical technique for determining the chemical composition of a sample based on spectral analysis of inelastically scattered photons. As a vibrational spectroscopy technique, Raman is sensitive to both the molecular composition and molecular structure of the sample being interrogated and is thus highly sensitive to chemical changes due to disease pathology.

In this thesis we have investigated several improvements aimed at developing a clinically viable intracoronary Raman system, including high wavenumber Raman, Raman line imaging, and optical modeling for improving catheter design. We investigated HW Raman and Raman line imaging through the creation of an *ex vivo* Raman database, consisting of FP and HW Raman line images with correlated histology for 69 aortic lesions. Rigorous data acquisition, validation, and documentation procedures were implemented to ensure that all specimens had high quality spectral and histological data and that the spectral acquisition sites were accurately registered with the histology. All data was captured electronically, and a graphical user interface was developed to enable easy visualization of all the data for

each specimen on a single screen, which ultimately aided in the lesion characterization data analysis. To date, 69 lesions have been acquired from nine subjects, mostly representing a range of lipid-rich lesions. Future work includes increasing the number of specimens and subjects in the database.

To analyze the database spectra, we developed a spectral model consisting of basis spectra representing the most important chemical components in the arterial wall. The model included free cholesterol, cholesterol ester, triglyceride, and two separate protein components. Additionally, the FP model contained β-carotene and calcification, both of which have no spectral features in the HW region. The model was fit to the database spectra and was found to accurately account for all the significant bands in the FP and HW Raman spectra. A subset of the model fit coefficients were retained for lesion characterization: free cholesterol, cholesterol ester, protein, and calcification (FP only).

A subset of the database spectra were selected to create a single site Raman database, which were histologically classified into calcified, non-lipid, and lipid-rich lesions. The model fit coefficients were used to diagnose these sites based on the FP and HW Raman spectra. By calculating and thresholding the lipid-to-protein ratio, defined as the sum of the free and esterified cholesterol coefficients divided by the sum of the protein coefficients, spectra from both the FP and HW Raman regions were capable of diagnosing lipid-rich lesions, with sensitivity/specificity of 88%/84% and 95%/85%, respectively.

While both the FP and HW regions are capable of diagnosing lipid-rich vs. non-lipid lesions, each region has several advantages and drawbacks. The HW spectra exhibit increased Raman signal over the FP spectra. Additionally, the HW region has an easier to remove background spectrum, due to both reduced fluorescence and the fact that many of the optical components that contribute to the background in the FP region have a negligible signal in the HW region. Furthermore, the HW spectra have fewer spectral features than the FP region, so fitting HW spectra to the spectral model is easier. However, the HW region has two significant drawbacks: it cannot be used to detect the presence of calcifications, and it is not highly sensitive to cholesterol vs. cholesterol esters.

In contrast, the fingerprint region, as its name implies, is sensitive to minor differences in molecular composition and structure, such as between cholesterol and cholesterol esters, and the FP region can be used to detect the presence of calcification. However, the increased sensitivity and increased number of Raman bands in the FP region also makes the

spectral model more sensitive to errors in determining the model basis spectra. Even worse, there are more possible sources of error in the FP region than in the HW region. First, basis spectra not acquired from samples *in situ* may exhibit spectral shifts due to environmental conditions, and because the FP spectra have sharp bands, these spectral shifts can potentially be extremely damaging to model accuracy. Further, comparison of relative FP band intensities relies on accurate background subtraction, which is more difficult in the FP region because of increased fluorescence and the possibility of contamination by Raman and fluorescence arising from optical components in the Raman system. Finally, FP spectra have less signal than HW spectra and thus lower SNR. Thus, while the FP region has more information than the HW region, it is also more prone to confounding effects that introduce errors.

The database data was also used to assess the utility of Raman line imaging. Raman line images provide a more complete characterization of heterogeneous lesions than single site spectra. By simple visual inspection of the spectral model fit coefficients as a function of position, it possible to determine the presence and size of an atherosclerotic lesion.

We then turned our attention to the development of a fiber optic Raman catheter. A significant technical challenge in FP Raman probe fabrication is the need for two distinct distal optical filters to avoid contaminating the tissue Raman signal with a large background arising from the silica fiber. Through the use of a patterned monolithic optical filter, we designed a 775 μm diameter Raman probe that is simpler to fabricate than previously demonstrated Raman probes.

In the most common intravascular catheter-based imaging paradigm, a sub-millimeter catheter is free-floating within the lumen of a $2-3$ mm diameter artery; thus the distance between the probe and the tissue surface is variable due to lateral catheter motion. For this reason, we assessed Raman signal collection as a function of distance between the probe and the tissue surface, with both saline and blood as the intervening medium. Due to high blood absorption, the probe was not capable of acquiring tissue Raman spectra through more than ~500 μm of blood, suggesting that successful intracoronary Raman spectroscopy will require the use of either a contact catheter or a saline flush to purge the blood prior to spectral acquisition. Tissue Raman spectra were reliably obtained through ~2 mm of saline.

The Raman probe was incorporated into a side-viewing Raman catheter by adding a

45° mirror and enclosing the assembly in an outer sheath, and this side-viewing Raman catheter was tested in a swine xenograft model. In this model, human coronary arteries obtained from cadaver are grafted onto a living swine's circulatory system, enabling testing of intracoronary devices for assessing real human pathology in an environment closely resembling human coronary physiology. The side-viewing Raman catheter was able to acquire interpretable Raman spectra with an integration time of 4 s, albeit with relatively low SNR. In the low SNR environment provided by the catheter, the technical advantages of the HW region stand out. The higher magnitude signal in the HW region combined with the relatively featureless background make HW data analysis simpler than in the FP region.

We developed a Raman Monte Carlo model for Raman scattering, and utilized it to develop a simulation technique for modeling catheter performance and optimizing catheter design. The modeling technique considers both propagation of the excitation beam and resulting generation of Raman scattered light, as well as the propagation of the returning Raman light. In our implementation, we have incorporated two software packages, a Monte Carlo package for accurate modeling of optical propagation in tissue, and a commercial ray tracing package for accurate modeling of propagation through complex optics. By simulating all aspects of Raman generation and propagation, the modeling technique can assess the sampling volume and collection efficiency of an arbitrary Raman probe.

We experimentally validated our simulation technique by measuring the sampling volume of a two-fiber Raman probe. A small diamond chip was treated as a Raman point source and immersed in a scattering phantom solution. The Raman probe then measured the magnitude of the diamond signal as the diamond was scanned through the probe's sampling volume. For two different phantom solutions with optical properties similar to human aorta, the experimental and simulation derived sampling volumes show good qualitative and quantitative agreement.

Future work includes using the validated Raman optical simulation technique to optimize Raman catheter design, through optimization metrics such as sampling volume and collection efficiency. Several alternate optical designs will be explored, including the use of custom distal optical lenses to increase collection efficiency. The simulation technique will also be used to assess probe performance through the range of environmental conditions expected for the probe, such as varying catheter-tissue distance.

The lesion characterization techniques described in this thesis can be expanded into

several interesting directions. For example, quantitative analysis of the spatial information in the Raman line images is expected to enhance lesion characterization and may even be able to identify markers for lesion vulnerability based on spatial changes in lesion composition. In addition, the spectral model can be refined and externally validated to enable more detailed spectral decomposition of lipids, particularly cholesterol and the individual cholesterol esters, which are thought to be important during lesion development.

The Raman line images in the *ex vivo* database form the beginnings of a platform for further development of characterization algorithms and for testing the existing hypotheses relating chemical composition to lesion morphology. However, more subjects need to be incorporated into the database to achieve statistically significant numbers for each lesion type and to address inter-patient variability.

In the longer term, an intracoronary Raman spectroscopy system has several potential applications. Raman spectroscopy is already capable of diagnosing calcified, non-lipid, and lipid-rich plaques *ex vivo*; a viable catheter-based Raman system will enable *in vivo* diagnosis. Furthermore, a viable *in vivo* Raman system will allow for longitudinal studies of atherosclerosis, enabling chemical mapping of the same subject at the same location, over multiple time points. Longitudinal studies for studying lesion progression in the context of chemical composition may aid in the understanding of the disease. The Raman system can also be used to assess long-term efficacy of treatments for atherosclerosis, such as lipid-lowering drug therapy and stents.

In summary, in this thesis we have investigated methods that may further the development of clinically viable intracoronary Raman spectroscopy. We have created a database of FP and HW Raman line images with correlated histology and used it to assess the relative merits of the FP and HW regions for lesion characterization. Both regions have similar diagnostic capability for detecting non-lipid vs. lipid-rich plaques, but only the FP region has sensitivity to calcification. In general, the FP region has higher molecular specificity than the HW region, but that same property makes FP spectra more difficult to analyze than HW spectra, especially in a low SNR regime. The HW region is simpler to work with, both in terms of data analysis and catheter design. Ultimately, the determination of which wavenumber region to use depends on whether the enhanced chemical information in the FP region is more clinically valuable than the ease-of-use offered by the HW region. In either case, a Raman catheter is needed, and to that end, we have developed a validated Raman

optical simulation technique to guide probe design. The union of the work presented in this thesis provide useful insights toward solving the challenges in developing a catheter-based intracoronary Raman system.

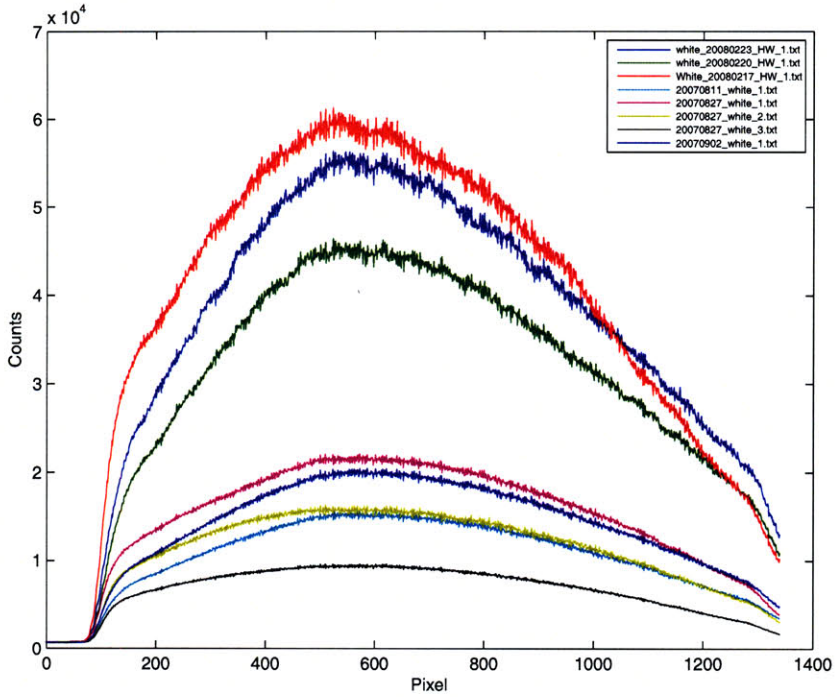
Appendix A

Photoresponse Non-Uniformity

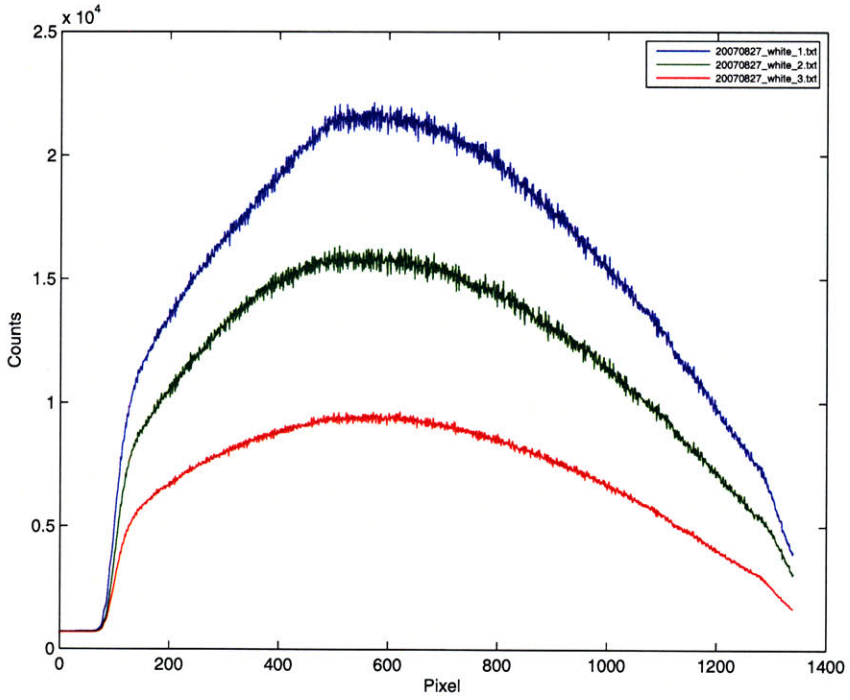
The raw spectral data exhibit a high frequency signal variation that can obscure narrow, low intensity Raman bands. Photoresponse non-uniformity, or pixel response non-uniformity (PRNU), arises from imperfections introduced during the CCD array fabrication process. As a result, the quantum efficiency of each pixel is nonuniform, so even when the CCD array is uniformly illuminated with a fixed-intensity signal, the measured signal is not uniform [125].

In general, a common method of correcting for PRNU is to perform flat-field normalization, *i.e.* divide each data image by a flat-field image acquired with the CCD illuminated by a uniform light source. In Raman spectroscopy, the analogous approach is to acquire the spectrum of a diffusely reflected white light source. A typical configuration it is to place packed barium sulfate ($BaSO_4$) powder, a reflectance standard that reflects uniformly over our wavelength range, at the system's focal plane and illuminate it with a tungsten white light source. An important feature of this configuration is that the illumination conditions closely match the experimental conditions of the returning Raman light. Incoming data spectra are divided by a normalized version of this white light spectrum, thus correcting for both PRNU and the system's spectral response simultaneously.

In practice, this procedure is not straightforward. The white light spectrum is highly dependent on the orientation of both the tungsten source and the $BaSO_4$ sample. Figure A-1 shows eight white light spectra taken on six different days. All eight spectra have different spectral features, even the three spectra taken on the same day (Figure A-1(b)). This is undesirable for a standard.



(a) Eight white light spectra



(b) Three white light spectra from a single day

Figure A-1: (a) White light spectra taken with the benchtop Raman system on different days. Three spectra from a single day are shown in (b); these were taken with the white light source held at varying distance and angle with respect to the $BaSO_4$, giving varying levels of overall intensity and noticeably different spectral shapes.

Simply dividing the data by a white light spectrum taken on the same day does not seem to be the best way to correct to for PRNU. While the white light spectrum is expected to have a smooth, low order polynomial shape, the measured white light spectra exhibit extra variation, with a period of approximately 50 pixels, throughout the spectrum (most readily apparent on three highest intensity spectra of Figure A-1(a)). Dividing the data by the white light spectrum note only removes PRNU, but also adds uncharacterized frequency dependent variation to the processed data.

Instead, we use the white light spectra to estimate PRNU correction factors, which we then use to remove PRNU from our data. For notational simplicity, we will refer to this correction factor as quantum efficiency, η.[1] We assume that our measured spectrum, $S_m(\omega)$, is related to the real spectrum of the light impinging on the CCD, $S_o(\omega)$ by

$$S_m(\omega) = S_o(\omega)\eta(\omega). \tag{A.1}$$

In the benchtop Raman system, each pixel corresponds to a particular frequency range, so the variable ω is also standing in for column location on the CCD. We wish to calculate

$$\eta(\omega) = \frac{S_m(\omega)}{S_o(\omega)}, \tag{A.2}$$

but $S_o(\omega)$ is unknown and thus must be estimated.

To determine $\hat{S}_o(\omega)$, the estimate of $S_o(\omega)$, we first subtract the dark spectrum from the white light spectrum to remove the CCD's fixed offset signal. Next, we estimate $S_o(\omega)$ by locally fitting a polynomial to the measured $S_m(\omega)$. We assume that $S_o(\omega)$ is smoothly varying, *i.e.* there are no sharp peaks in the spectrum. We choose a small moving window ($\sim$ 100 pixel length) of the measured spectrum, $S_m(\omega)$, and fit a fifth order polynomial to estimate $S_o(\omega)$ within each window. $\hat{S}_o(\omega)$ is constructed by selecting the central portion of the polynomial fit for each window, which avoids edge effects in the polynomial fits (Figure A-2).

Finally, we calculate

$$\hat{\eta}(\omega) = \frac{S_m(\omega)}{\hat{S}_o(\omega)} \tag{A.3}$$

In Figure A-3, we have estimated $\hat{\eta}(\omega)$ for all eight white light spectra shown in Figure A-1.

[1]Technically, we are estimating a quantity presumed to be proportional to the quantum efficiency, not the actual quantum efficiency

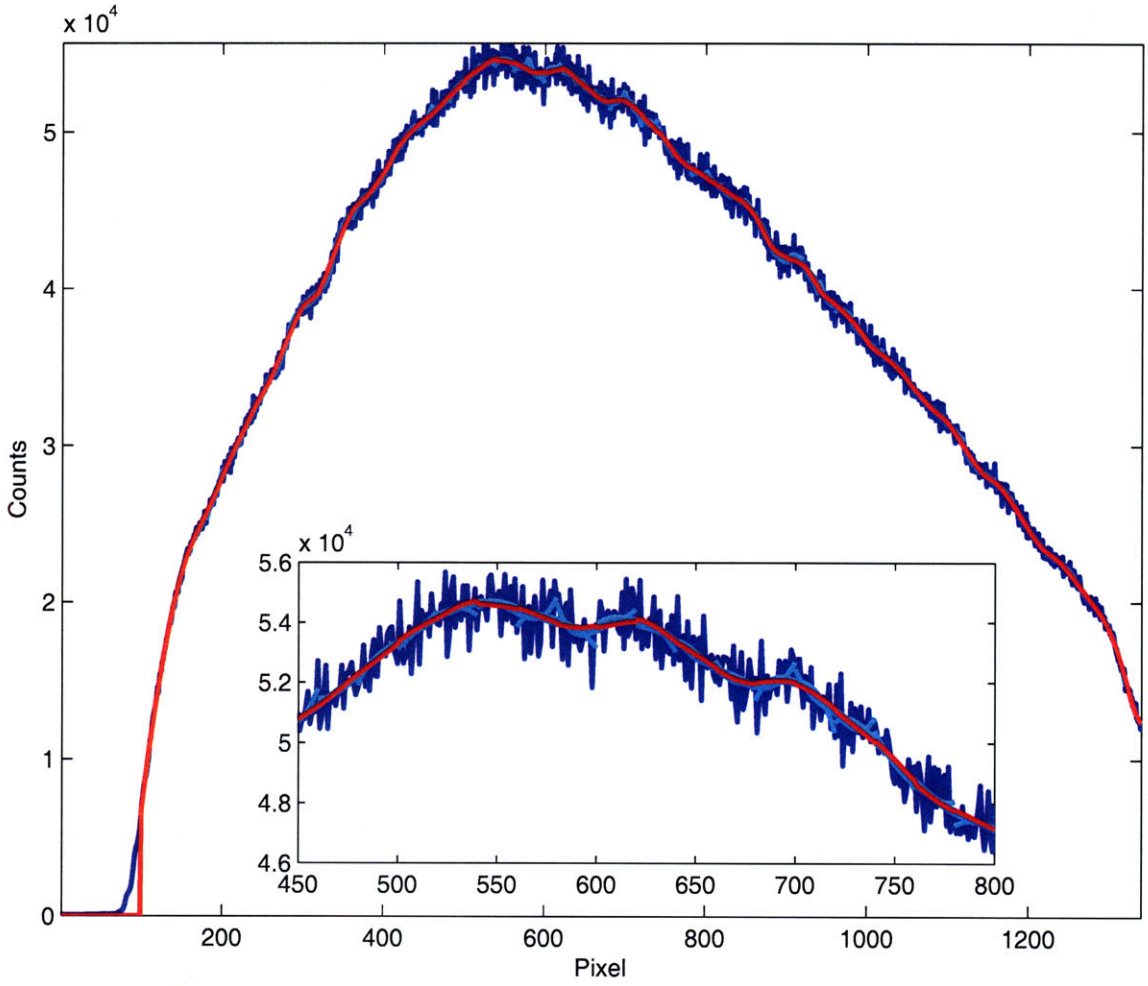
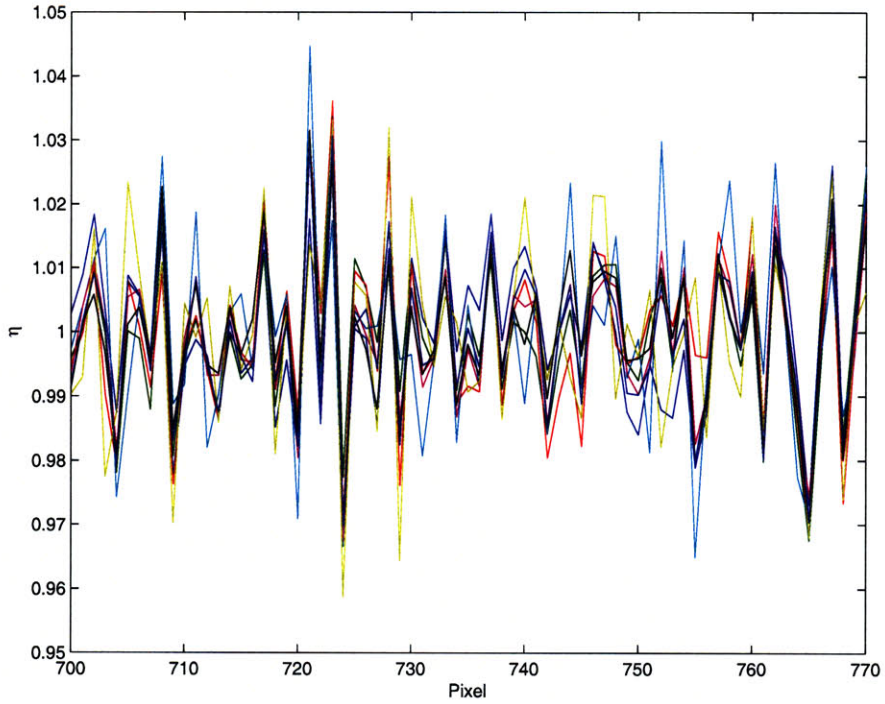


Figure A-2: An example estimate of $\hat{S}_o(\omega)$. The original white light data is shown in blue. The central portion of the windowed polynomial fits (cyan) were selected to estimate $\hat{S}_o(\omega)$ (red). Overlapping windows were necessary to avoid edge effects in the windowed polynomial fits.

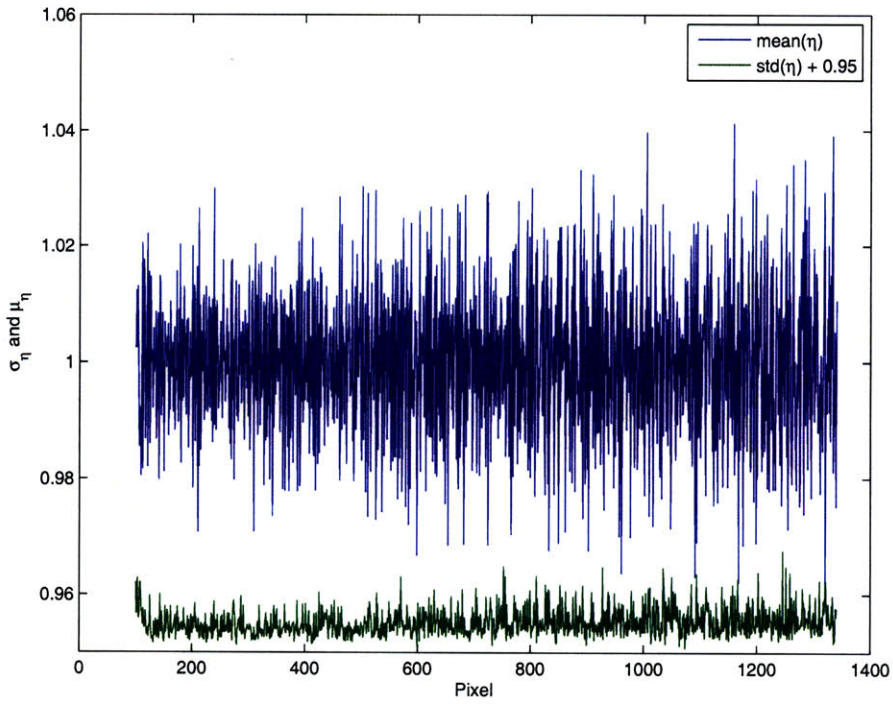
The calculated $\hat{\eta}$ values are similar for all eight white light spectra, and $\hat{\eta}$ ranges from 0.96 to 1.04 for all pixels. The standard deviation of $\hat{\eta}$ for the eight spectra is less than 0.01 for all pixels. Figure A-3 depicts the $\hat{\eta}$ that was used for the database data.

Despite these differences in quantum efficiency being subtle, they have significant impact on the data. Some examples of corrected database data are shown in Figures A-4 and A-5. For display, the PRNU-corrected spectra have also been filtered with a Savitkzy-Golay filter [102, 126] with a polynomial order of 2 and frame size of 11. The filter is built into MATLAB's Signal Processing Toolbox (`sgolayfilt`).

The calculated PRNU correction factor is specific to the CCD used to acquire the data. Furthermore, since spectra are typically acquired by vertically binning some or all of the

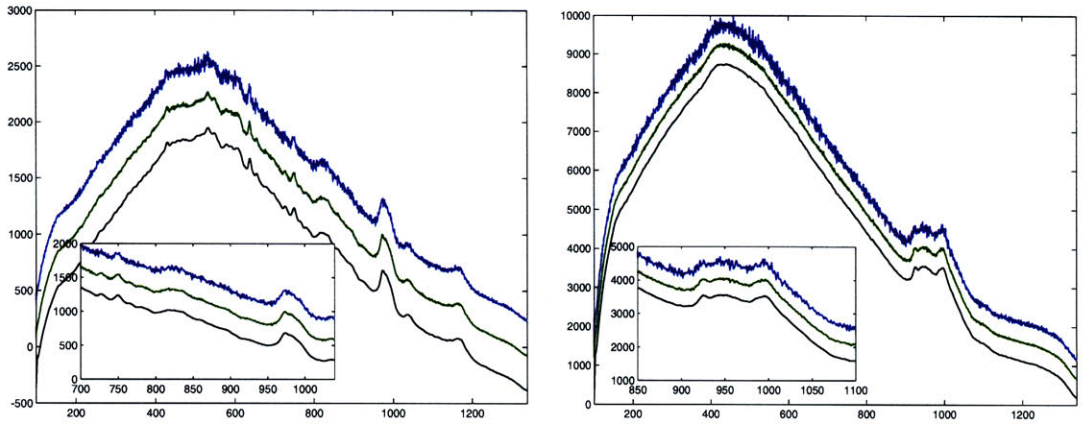


(a) $\hat{\eta}$ for all spectra



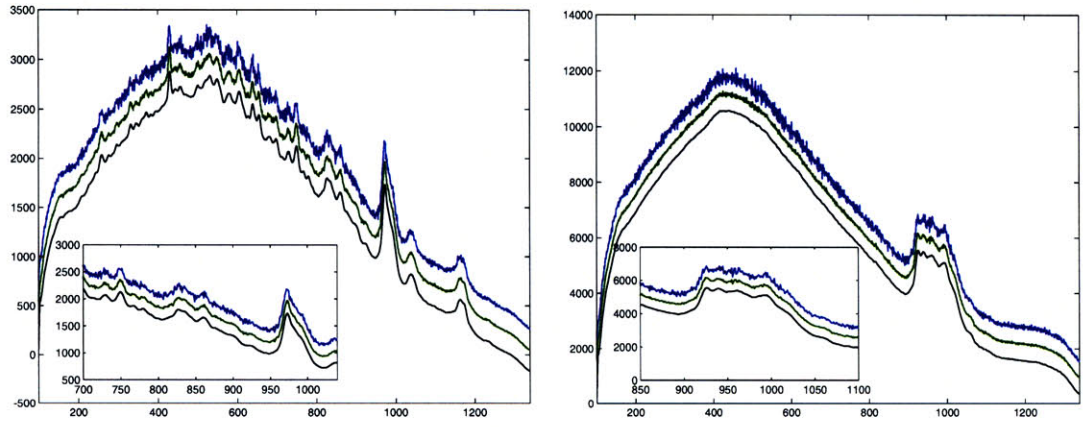
(b) Mean and standard deviation of $\hat{\eta}$

Figure A-3: $\hat{\eta}$ for the database data. In (a), colored lines represent $\hat{\eta}$ calculated for same color white light spectrum in Figure A-1(a). In (b), the mean $\hat{\eta}$ value for the eight white light spectra is shown, as well as the standard deviation, offset by 0.95.



(a) Fingerprint spectrum **(b)** High wavenumber spectrum

Figure A-4: Example of PRNU correction on database data (specimen UMD019). The original data (blue), after PRNU correction (green), and after Savitzky-Golay filtering (black). All spectra are offset for clarity.



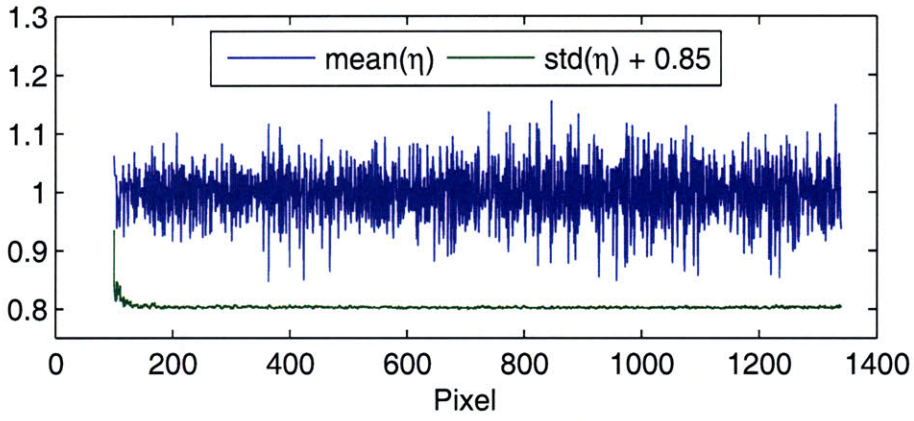
(a) Fingerprint spectrum **(b)** High wavenumber spectrum

Figure A-5: Example of PRNU correction on database data (UMD020). The original data (blue), after PRNU correction (green), and after Savitzky-Golay filtering (black). All spectra are offset for clarity.

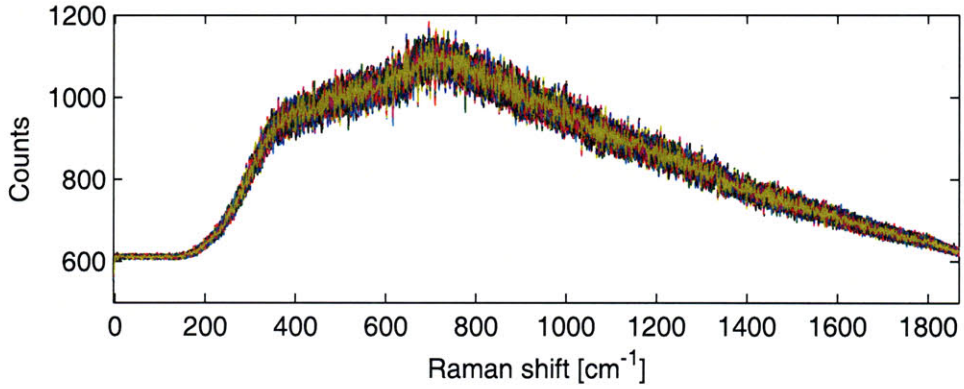
rows on the CCD, the correction factor is dependent on the illumination conditions. Even if all rows are included in the binned measurement, illuminating the top half of the chip will result in a different $\hat{\eta}$ than illuminating the bottom half, because the quantum efficiencies of different pixels are being incorporated into the measurements.

It is not necessary to use white light spectra to calculate PRNU correction factors; any spectrum with either a known spectral shape or a smoothly varying spectral shape that can be estimated is suitable for estimating $\hat{\eta}$.

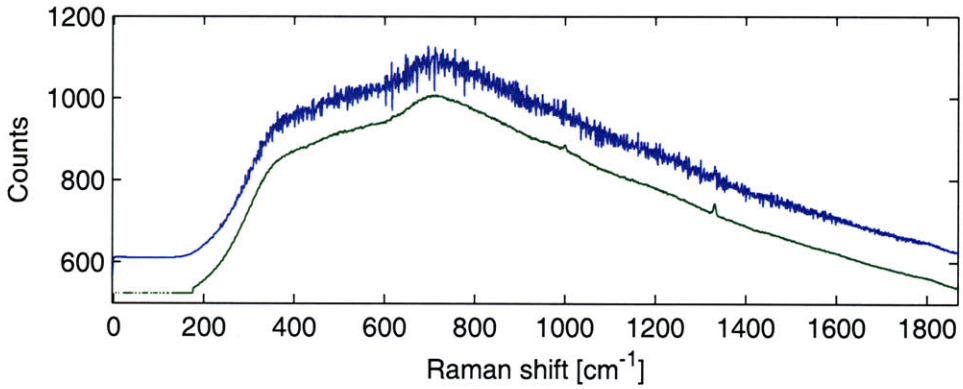
The Raman probe sampling volume experiments in Chapter 6 had different illumination conditions than the database data, so a new $\hat{\eta}$ was estimated (Figure A-6(a)). The increased variation in $\hat{\eta}$ across pixels, compared to the database correction factor (Figure A-3), occurs because the probe illuminates fewer rows of the CCD. PRNU correction was vital in this case. In one spectrum from the sampling volume experiments, two small Raman peaks that were not discernible in the frame averaged data become readily apparent in the PRNU corrected data (Figure A-6(c)).



(a) $\hat{\eta}$



(b) Raw data (80 mW, 230 frames @ 0.25 s/frame



(c) Mean of raw data (blue) and PRNU corrected data (green)

Figure A-6: PRNU correction on sampling volume data. (a) Correction factor for the sampling volume experiments of Chapter 6. (b) Raw data for an example sampling volume spectrum. The small peaks at 1001 cm^{-1} (polystyrene) and 1332 cm^{-1} (diamond) become readily apparent in the preprocessed data in (c). No additional smoothing was performed.

Bibliography

[1] D. Lloyd-Jones, R. Adams, M. Carnethon, G. De Simone, T. B. Ferguson, K. Flegal, E. Ford, K. Furie, A. Go, K. Greenlund, N. Haase, S. Hailpern, M. Ho, V. Howard, B. Kissela, S. Kittner, D. Lackland, L. Lisabeth, A. Marelli, M. McDermott, J. Meigs, D. Mozaffarian, G. Nichol, C. O'Donnell, V. Roger, W. Rosamond, R. Sacco, P. Sorlie, R. Stafford, J. Steinberger, T. Thom, S. Wasserthiel-Smoller, N. Wong, J. Wylie-Rosett, and Y. Hong. Heart Disease and Stroke Statistics–2009 Update: A report from the American Heart Association Statistics Committee and Stroke Statistics Subcommittee. *Circulation*, 119(3):E21–181, 2009.

[2] R. Ross. Atherosclerosis — An inflammatory disease. *The New England Journal of Medicine*, 340(2):115–126, 1999.

[3] P. Libby. Inflammation in atherosclerosis. *Nature*, 420(6917):868–874, 2002.

[4] P. Libby. Inflammation and cardiovascular disease mechanisms. *American Journal of Clinical Nutrition*, 83(2):456S–460, 2006.

[5] H. Watkins and M. Farrall. Genetic susceptibility to coronary artery disease: from promise to progress. *Nature Reviews Genetics*, 7:163–173, 2006.

[6] H. C. Stary, A. B. Chandler, S. Glagov, J. R. Guyton, W. Insull, Jr., M. E. Rosenfeld, S. A. Schaffer, C. J. Schwartz, W. D. Wagner, and R. W. Wissler. A definition of initial, fatty streak, and intermediate lesions of atherosclerosis. A report from the Committee on Vascular Lesions of the Council on Arteriosclerosis, American Heart Association. *Circulation*, 89(5):2462–2478, 1994.

[7] C. K. Glass and J. L. Witztum. Atherosclerosis: The road ahead. *Cell*, 104(4):503–516, 2001.

[8] K. Yokoya, H. Takatsu, T. Suzuki, H. Hosokawa, S. Ojio, T. Matsubara, T. Tanaka, S. Watanabe, N. Morita, K. Nishigaki, G. Takemura, T. Noda, S. Minatoguchi, and H. Fujiwara. Process of progression of coronary artery lesions from mild or moderate stenosis to moderate or severe stenosis: A study based on four serial coronary arteriograms per year. *Circulation*, 100(9):903–909, 1999.

[9] R. Virmani, A. P. Burke, A. Farb, and F. D. Kolodgie. Pathology of the unstable plaque. *Progress in Cardiovascular Diseases*, 44(5):349–356, 2002.

[10] R. Virmani, A. P. Burke, A. Farb, and F. D. Kolodgie. Pathology of the vulnerable plaque. *Journal of the American College of Cardiology*, 47(C):C13–18, 2006.

[11] R. Virmani, F. D. Kolodgie, A. P. Burke, A. Farb, and S. M. Schwartz. Lessons from sudden coronary death: A comprehensive morphological classification scheme for atherosclerotic lesions. *Arteriosclerosis, Thrombosis, and Vascular Biolology*, 20(5):1262–1275, 2000.

[12] F. D. Kolodgie, A. P. Burke, A. Farb, H. K. Gold, J. Yuan, J. Narula, A. V. Finn, and R. Virmani. The thin-cap fibroatheroma: a type of vulnerable plaque: the major precursor lesion to acute coronary syndromes. *Current Opinion in Cardiology*, 16(5):285–92, 2001.

[13] A. C. Van der Wal, A. E. Becker, C. M. Vanderloos, and P. K. Das. Site of intimal rupture or erosion of thrombosed coronary atherosclerotic plaques is characterized by an inflammatory process irrespective of the dominant plaque morphology. *Circulation*, 89(1):36–44, 1994.

[14] A. Farb, A. P. Burke, A. L. Tang, T. Y. Liang, P. Mannan, J. Smialek, and R. Virmani. Coronary plaque erosion without rupture into a lipid core. A frequent cause of coronary thrombosis in sudden coronary death. *Circulation*, 93(7):1354–63, 1996.

[15] F. D. Kolodgie, A. P. Burke, T. N. Wight, and R. Virmani. The accumulation of specific types of proteoglycans in eroded plaques: a role in coronary thrombosis in the absence of rupture. *Current Opinion in Lipidology*, 15(5):575–82, 2004.

[16] J. A. Schaar, J. E. Muller, E. Falk, R. Virmani, V. Fuster, P. W. Serruys, A. Colombo, C. Stefanadis, S. Ward Casscells, P. R. Moreno, A. Maseri, and A. F. W. van der Steen. Terminology for high-risk and vulnerable coronary artery plaques. *European Heart Journal*, 25(12):1077–1082, 2004.

[17] E. Arbustini, M. Grasso, M. Diegoli, A. Pucci, M. Bramerio, D. Ardissino, L. Angoli, S. Deservi, E. Bramucci, A. Mussini, G. Minzioni, M. Vigano, and G. Specchia. Coronary atherosclerotic plaques with and without thrombus in ischemic heart syndromes. A morphological, immunohistochemical, and biochemical study. *American Journal of Cardiology*, 68(7):B36–B50, 1991.

[18] G. Rioufol, G. Finet, I. Ginon, X. Andre-Fouet, R. Rossi, E. Vialle, E. Desjoyaux, G. Convert, J. F. Huret, and A. Tabib. Multiple atherosclerotic plaque rupture in acute coronary syndrome: a three-vessel intravascular ultrasound study. *Circulation*, 106(7):804–8., 2002.

[19] L. M. Biasucci, G. Liuzzo, R. L. Grillo, G. Caligiuri, A. G. Rebuzzi, A. Buffon, F. Summaria, F. Ginnetti, G. Fadda, and A. Maseri. Elevated levels of C-reactive protein at discharge in patients with unstable angina predict recurrent instability. *Circulation*, 99(7):855–60, 1999.

[20] A. Maseri and V. Fuster. Is there a vulnerable plaque? *Circulation*, 107(16):2068–71, 2003.

[21] M. Naghavi, P. Libby, E. Falk, S. W. Casscells, S. Litovsky, J. Rumberger, J. J. Badimon, C. Stefanadis, P. Moreno, G. Pasterkamp, Z. Fayad, P. H. Stone, S. Waxman, P. Raggi, M. Madjid, A. Zarrabi, A. Burke, C. Yuan, P. J. Fitzgerald, D. S. Siscovick, C. L. de Korte, M. Aikawa, K. E. J. Airaksinen, G. Assmann, C. R. Becker, J. H.

Chesebro, A. Farb, Z. S. Galis, C. Jackson, I. K. Jang, W. Koenig, R. A. Lodder, K. March, J. Demirovic, M. Navab, S. G. Priori, M. D. Rekhter, R. Bahr, S. M. Grundy, R. Mehran, A. Colombo, E. Boerwinkle, C. Ballantyne, W. Insull, R. S. Schwartz, R. Vogel, P. W. Serruys, G. K. Hansson, D. P. Faxon, S. Kaul, H. Drexler, P. Greenland, J. E. Muller, R. Virmani, P. M. Ridker, D. P. Zipes, P. K. Shah, and J. T. Willerson. From vulnerable plaque to vulnerable patient - A call for new definitions and risk assessment strategies: Part II. *Circulation*, 108(15):1772–1778, 2003.

[22] A. Buffon, L. M. Biasucci, G. Liuzzo, G. D'Onofrio, F. Crea, and A. Maseri. Widespread coronary inflammation in unstable angina. *The New England Journal of Medicine*, 347(1):5–12, 2002.

[23] P. M. Ridker, N. Rifai, L. Rose, J. E. Buring, and N. R. Cook. Comparison of C-reactive protein and low-density lipoprotein cholesterol levels in the prediction of first cardiovascular events. *The New England Journal of Medicine*, 347(20):1557–1565, 2002.

[24] C. V. Felton, D. Crook, M. J. Davies, and M. F. Oliver. Relation of plaque lipid composition and morphology to the stability of human aortic plaques. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 17(7):1337–1345, 1997.

[25] P. D. Richardson. Biomechanics of plaque rupture: Progress, problems, and new frontiers. *Annals of Biomedical Engineering*, 30(4):524–536, 2002.

[26] M. A. Pulido, D. J. Angiolillo, and M. A. Costa. Imaging of atherosclerotic plaque. *International Journal of Cardiovascular Imaging*, 20(6):553–559, 2004.

[27] A. Taylor, L. J. Shaw, Z. Fayad, D. O'Leary, B. G. Brown, S. Nissen, D. Rader, and P. Raggi. Tracking atherosclerosis regression: a clinical tool in preventive cardiology. *Atherosclerosis*, 180(1):1–10, 2005.

[28] C. R. Becker. Noninvasive assessment of coronary atherosclerosis by multidetector-row computed tomography. *Expert Review of Cardiovascular Therapy*, 2(5):721–7, 2004.

[29] F. Knollmann, F. Ducke, L. Krist, T. Kertesz, R. Meyer, H. Guski, and R. Felix. Quantification of atherosclerotic coronary plaque components by submillimeter computed tomography. *International Journal of Cardiovascular Imaging*, 2007.

[30] S. Motoyama, T. Kondo, M. Sarai, A. Sugiura, H. Harigaya, T. Sato, K. Inoue, M. Okumura, J. Ishii, H. Anno, R. Virmani, Y. Ozaki, H. Hishida, and J. Narula. Multislice computed tomographic characteristics of coronary lesions in acute coronary syndromes. *Journal of the American College of Cardiology*, 50(4):319–26, 2007.

[31] Z. Fayad and V. Fuster. The human high-risk plaque and its detection by magnetic resonance imaging. *American Journal of Cardiology*, 88(2-A):42E–45E, 2001.

[32] C. Yuan, S. X. Zhang, N. L. Polissar, D. Echelard, G. Ortiz, J. W. Davis, E. Ellington, M. S. Ferguson, and T. S. Hatsukami. Identification of fibrous cap rupture with magnetic resonance imaging is highly associated with recent transient ischemic attack or stroke. *Circulation*, 105(2):181–5, 2002.

[33] M. Takano, S. Inami, F. Ishibashi, K. Okamatsu, K. Seimiya, T. Ohba, S. Sakai, and K. Mizuno. Angioscopic follow-up study of coronary ruptured plaques in nonculprit lesions. *Journal of the American College of Cardiology*, 45(5):652–658, 2005.

[34] J. A. Schaar, C. L. de Korte, F. Mastik, C. Strijder, G. Pasterkamp, E. Boersma, P. W. Serruys, and A. F. W. van der Steen. Characterizing vulnerable plaque features with intravascular elastography. *Circulation*, 108(21):2636–2641, 2003.

[35] C. Stenfanadis, L. Diamantopoulos, C. Vlachopoulos, E. Tsiamis, J. Dernellis, K. Toutouzas, E. Stefandi, and P. Toutouzas. Thermal heterogeneity within human atherosclerotic coronary arteries detected in vivo: A new method of detection by application of a special thermography catheter. *Circulation*, 99:1965–1971, 1999.

[36] T. Hiro, C. Y. Leung, S. De Guzman, V. J. Caiozzo, A. R. Farvid, H. Karimi, R. H. Helfant, and J. M. Tobis. Are soft echoes really soft? Intravascular ultrasound assessment of mechanical properties in human atherosclerotic tissue. *American Heart Journal*, 133(1):1–7, 1997.

[37] T. Hagenaars, E. J. Gussenhoven, S. E. Kranendonk, J. D. Blankensteijn, J. Honkoop, E. van der Linden, and A. van der Lugt. Early experience with intravascular ultrasound in evaluating the effect of statins on femoropopliteal arterial disease: hypothesis-generating observations in humans. *Cardiovascular Drugs and Therapy*, 14(6):635–41, 2000.

[38] S. E. Nissen and P. Yock. Intravascular ultrasound: novel pathophysiological insights and current clinical applications. *Circulation*, 103(4):604–16, 2001.

[39] T. Okimoto, M. Imazu, Y. Hayashi, H. Fujiwara, H. Ueda, and N. Kohno. Atherosclerotic plaque characterization by quantitative analysis using intravascular ultrasound: correlation with histological and immunohistochemical findings. *Circulation Journal*, 66(2):173–7., 2002.

[40] K. Nasu, E. Tsuchikane, O. Katoh, D. G. Vince, R. Virmani, J. F. Surmely, A. Murata, Y. Takeda, T. Ito, M. Ehara, T. Matsubara, M. Terashima, and T. Suzuki. Accuracy of in vivo coronary plaque morphology assessment: a validation study of in vivo virtual histology compared with in vitro histopathology. *Journal of the American College of Cardiology*, 47(12):2405–12, 2006.

[41] S. Flacke, S. Fischer, M. J. Scott, R. J. Fuhrhop, J. S. Allen, M. McLean, P. Winter, G. A. Sicard, P. J. Gaffney, S. A. Wickline, and G. M. Lanza. Novel MRI contrast agent for molecular imaging of fibrin: implications for detecting vulnerable plaques. *Circulation*, 104(11):1280–5, 2001.

[42] J. Chen, C. H. Tung, U. Mahmood, V. Ntziachristos, R. Gyurko, M. C. Fishman, P. L. Huang, and R. Weissleder. In vivo imaging of proteolytic activity in atherosclerosis. *Circulation*, 105(23):2766–71, 2002.

[43] D. E. Sosnovik, E. A. Schellenberger, M. Nahrendorf, M. S. Novikov, T. Matsui, G. Dai, F. Reynolds, L. Grazette, A. Rosenzweig, R. Weissleder, and L. Josephson. Magnetic resonance imaging of cardiomyocyte apoptosis with a novel magneto-optical nanoparticle. *Magnetic Resonance in Medicine*, 54(3):718–24, 2005.

[44] K. C. Briley-Saebo, W. J. Mulder, V. Mani, F. Hyafil, V. Amirbekian, J. G. Aguinaldo, E. A. Fisher, and Z. A. Fayad. Magnetic resonance imaging of vulnerable atherosclerotic plaques: current imaging strategies and molecular imaging probes. *Journal of Magnetic Resonance Imaging*, 26(3):460–79, 2007.

[45] S. M. O'Malley, M. Vavuranakis, M. Naghavi, and I. A. Kakadiaris. Intravascular ultrasound-based imaging of vasa vasorum for the detection of vulnerable atherosclerotic plaque. *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2005*, 8(Pt 1):343–51, 2005.

[46] M. Vavuranakis, I. A. Kakadiaris, T. G. Papaioannou, S. M. O'Malley, S. Carlier, M. Naghavi, and C. Stefanadis. Contrast-enhanced intravascular ultrasound: combining morphology with activity-based assessment of plaque vulnerability. *Expert Review of Cardiovascular Therapy*, 5(5):917–25, 2007.

[47] H. Yabushita, B. E. Bouma, S. L. Houser, T. Aretz, I. K. Jang, K. H. Schlendorf, C. R. Kauffman, M. Shishkov, D. H. Kang, E. F. Halpern, and G. J. Tearney. Characterization of human atherosclerosis by optical coherence tomography. *Circulation*, 106(13):1640–1645, 2002.

[48] A. F. Low, G. J. Tearney, B. E. Bouma, and I. K. Jang. Technology insight: optical coherence tomography - current status and future development. *Nature Clinical Practice Cardiovascular Medicine*, 3(3):154–162, 2006.

[49] G. J. Tearney, H. Yabushita, S. L. Houser, H. T. Aretz, I. K. Jang, K. H. Schlendorf, C. R. Kauffman, M. Shishkov, E. F. Halpern, and B. E. Bouma. Quantification of macrophage content in atherosclerotic plaques by optical coherence tomography. *Circulation*, 107(1):113–119, 2003.

[50] I. K. Jang, B. E. Bouma, D. H. Kang, S. J. Park, S. W. Park, K. B. Seung, K. B. Choi, M. Shishkov, K. Schlendorf, E. Pomerantsev, S. L. Houser, H. T. Aretz, and G. J. Tearney. Visualization of coronary atherosclerotic plaques in patients using optical coherence tomography: comparison with intravascular ultrasound. *Journal of the American College of Cardiology*, 39(4):604–9, 2002.

[51] B. D. MacNeill, I. K. Jang, B. E. Bouma, N. Iftimia, M. Takano, H. Yabushita, M. Shishkov, C. R. Kauffman, S. L. Houser, H. T. Aretz, D. DeJoseph, E. F. Halpern, and G. J. Tearney. Focal and multi-focal plaque macrophage distributions in patients with acute and stable presentations of coronary artery disease. *Journal of the American College of Cardiology*, 44(5):972–9, 2004.

[52] I. K. Jang, G. J. Tearney, B. MacNeill, M. Takano, F. Moselewski, N. Iftima, M. Shishkov, S. Houser, H. T. Aretz, E. F. Halpern, and B. E. Bouma. In vivo characterization of coronary atherosclerotic plaque by use of optical coherence tomography. *Circulation*, 111(12):1551–5, 2005.

[53] L. Josephson, M. F. Kircher, U. Mahmood, Y. Tang, and R. Weissleder. Near-infrared fluorescent nanoparticles as combined MR/optical imaging probes. *Bioconjugate Chemistry*, 13(3):554–60, 2002.

[54] F. A. Jaffer, P. Libby, and R. Weissleder. Molecular imaging of cardiovascular disease. *Circulation*, 116(9):1052–61, 2007.

[55] S. W. E. van de Poll, J. T. Motz, J. R. Kramer, and M. S. Feld. Prospects of laser spectroscopy to detect vulnerable plaque. In D. Brown, editor, *Cardiovascular Plaque Rupture*, pages 281–303. Marcel Dekker, New York, 2002.

[56] F. Bosshart, U. Utzinger, O. M. Hess, J. Wyser, A. Mueller, J. Schneider, P. Niederer, M. Anliker, and H. P. Krayenbuehl. Fluorescence spectroscopy for identification of atherosclerotic tissue. *Cardiovascular Research*, 26(6):620–5, 1992.

[57] L. Marcu, M. Fishbein, J. Maarek, and W. Grundfest. Discrimination of human coronary artery atherosclerotic lipid-rich lesions by time-resolved laser-induced fluorescence spectroscopy. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 21(7):1244–50, 2001.

[58] K. Arakawa, K. Isoda, T. Ito, K. Nakajima, T. Shibuya, and F. Ohsuzu. Fluorescence analysis of biochemical constituents identifies atherosclerotic plaque with a thin fibrous cap. *Arteriosclerosis, Thrombosis, and Vascular Biology*, 22(6):1002–7, 2002.

[59] L. Marcu, Q. Fang, J. A. Jo, T. Papaioannou, A. Dorafshar, T. Reil, J. H. Qiao, J. D. Baker, J. A. Freischlag, and M. C. Fishbein. In vivo detection of macrophages in a rabbit atherosclerotic model by time-resolved laser-induced fluorescence spectroscopy. *Atherosclerosis*, 181(2):295–303, 2005.

[60] J. A. Jo, Q. Fang, T. Papaioannou, J. D. Baker, A. H. Dorafshar, T. Reil, J. H. Qiao, M. C. Fishbein, J. A. Freischlag, and L. Marcu. Laguerre-based method for analysis of time-resolved fluorescence data: application to in-vivo characterization and diagnosis of atherosclerotic lesions. *Journal of Biomedical Optics*, 11(2):021004, 2006.

[61] J. Wang, Y. J. Geng, B. Guo, T. Klima, B. N. Lal, J. T. Willerson, and W. Casscells. Near-infrared spectroscopic characterization of human advanced atherosclerotic plaques. *Journal of the American College of Cardiology*, 39(8):1305–13, 2002.

[62] K. Mizuno, A. Miyamoto, K. Satomura, A. Kurita, T. Arai, M. Sakurada, S. Yanagida, and H. Nakamura. Angioscopic coronary macromorphology in patients with acute coronary disorders. *Lancet*, 337(8745):809–12, 1991.

[63] Y. Uchida, F. Nakamura, T. Tomaru, T. Morita, T. Oshima, T. Sasaki, S. Morizuki, and J. Hirose. Prediction of acute coronary syndromes by percutaneous coronary angioscopy in patients with stable angina. *American Heart Journal*, 30(2):195–203, 1995.

[64] T. Ohtani, Y. Ueda, M. Shimizu, N. Kondou, A. Hirayama, and K. Kodama. Risk of ischemic events can be predicted by angioscopic evaluation of yellow plaques. *Supplement to Journal of the American College of Cardiology*, 39(5, Suppl. A), 2002.

[65] J. D. Caplan, S. Waxman, R. W. Nesto, and J. E. Muller. Near-infrared spectroscopy for the detection of vulnerable coronary artery plaques. *Journal of the American College of Cardiology*, 47(8 Suppl):C92–6, 2006.

[66] C. Raman and K. Krishnan. A new type of secondary radiation. *Nature*, 121:501–502, 1928.

[67] J. J. Baraga, M. S. Feld, and R. P. Rava. Rapid near-Infrared Raman spectroscopy of human tissue with a spectrograph and CCD detector. *Applied Spectroscopy*, 46(2):187–190, 1992.

[68] J. F. Brennan, III, T. J. Römer, R. S. Lees, A. M. Tercyak, J. R. Kramer, Jr., and M. S. Feld. Determination of human coronary artery composition by Raman spectroscopy. *Circulation*, 96(1):99–105, 1997.

[69] H. P. Buschman, J. T. Motz, G. Deinum, T. J. Römer, M. Fitzmaurice, J. R. Kramer, A. van der Laarse, A. V. Bruschke, and M. S. Feld. Diagnosis of human coronary atherosclerosis by morphology-based Raman spectroscopy. *Cardiovascular Pathology*, 10(2):59–68, 2001.

[70] T. J. Römer, J. F. B. Brennan, M. Fitzmaurice, M. L. Feldstein, G. Deinum, J. L. Myles, J. R. Kramer, R. S. Lees, and M. S. Feld. Histopathology of human coronary atherosclerosis by quantifying its chemical composition with Raman spectroscopy. *Circulation*, 97(9):878–885, 1998.

[71] J. J. Baraga. *In Situ Chemical Analysis of Biological Tissue: Vibrational Raman Spectroscopy of Human Atherosclerosis* . PhD thesis, Massachusetts Institute of Technology, 1992.

[72] R. L. McCreery. *Raman spectroscopy for chemical analysis.* John Wiley and Sons, New York, 2000.

[73] C. Krafft, G. Steiner, C. Beleites, and R. Salzer. Disease recognition by infrared and Raman spectroscopy. *Journal of Biophotonics*, 2(1):13–28, 2009.

[74] C. A. Lieber, S. K. Majumder, D. Billheimer, D. L. Ellis, and A. Mahadevan-Jansen. Raman microspectroscopy for skin cancer detection in vitro. *Journal of Biomedical Optics*, 13(2):024013, 2008.

[75] S. K. Majumder, M. D. Keller, F. I. Boulos, M. C. Kelley, and A. Mahadevan-Jansen. Comparison of autofluorescence, diffuse reflectance, and Raman spectroscopy for breast tissue discrimination. *Journal of Biomedical Optics*, 13(5):11, 2008.

[76] A. Robichaux-Viehoever, E. Kanter, H. Shappell, D. Billheimer, H. Jones, and A. Mahadevan-Jansen. Characterization of Raman spectra measured in vivo for the detection of cervical dysplasia. *Applied Spectroscopy*, 61(9):986–993, 2007.

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

[78] C. Kendall, N. Stone, N. Shepherd, K. Geboes, B. Warren, R. Bennett, and H. Barr. Raman spectroscopy, a potential tool for the objective identification and classification of neoplasia in Barrett's oesophagus. *Journal of Pathology*, 200(5):602–609, 2003.

[79] P. Crow, A. Molckovsky, N. Stone, J. Uff, B. Wilson, and L. M. Wongkeesong. Assessment of fiberoptic near-infrared Raman spectroscopy for diagnosis of bladder and prostate cancer. *Urology*, 65(6):1126–1130, 2005.

[80] C. Krafft, M. Kirsch, C. Beleites, G. Schackert, and R. Salzer. Methodology for fiber-optic Raman mapping and FTIR imaging of metastases in mouse brains. *Analytical and Bioanalytical Chemistry*, 389(4):1133–1142, 2007.

[81] G. Deinum, D. Rodrigues, T. J. Römer, M. Fitzmaurice, J. R. Kramer, and M. S. Feld. Histological classification of Raman spectra of human coronary artery atherosclerosis using principal component analysis. *Applied Spectroscopy*, 53(8):938–942, 1999.

[82] J. Salenius, J. F. B. Brennan, III, A. Miller, Y. Wang, T. Aretz, B. Sacks, R. Dasari, and M. Feld. Biochemical composition of human peripheral arteries examined with near-infrared Raman spectroscopy. *Journal of Vascular Surgery*, 27(4):710–9, 1998.

[83] J. T. Motz, M. Fitzmaurice, A. Miller, S. J. Gandhi, A. S. Haka, L. H. Galindo, R. R. Dasari, J. R. Kramer, and M. S. Feld. In vivo Raman spectral pathology of human atherosclerosis and vulnerable plaque. *Journal of Biomedical Optics*, 11(2):021003, 2006.

[84] L. Silveira, Jr., S. Sathaiah, R. A. Zngaro, M. T. T. Pacheco, M. C. Chavantes, and C. A. G. Pasqualucci. Correlation between near-infrared Raman spectroscopy and the histopathological analysis of atherosclerosis in human coronary arteries. *Lasers in Surgery and Medicine*, 30(4):290–297, 2002.

[85] G. V. Nogueira, L. Silveira, A. A. Martin, R. A. Zangaro, M. T. T. Pacheco, M. C. Chavantes, and C. A. Pasqualucci. Raman spectroscopy study of atherosclerosis in human carotid artery. *Journal of Biomedical Optics*, 10(3):031117, 2005.

[86] J. T. Motz, M. Hunter, L. H. Galindo, J. A. Gardecki, J. R. Kramer, R. R. Dasari, and M. S. Feld. Optical fiber probe for biomedical Raman spectroscopy. *Applied Optics*, 43(3):542–554, 2004.

[87] H. P. Buschman, E. T. Marple, M. L. Wach, B. Bennett, T. C. B. Schut, H. A. Bruining, A. V. Bruschke, A. van der Laarse, and G. J. Puppels. In vivo determination of the molecular composition of artery wall by intravascular Raman spectroscopy. *Analytical Chemistry*, 72(16):3771–3775, 2000.

[88] S. Koljenovic, T. C. Bakker Schut, R. Wolthuis, B. de Jong, L. Santos, P. J. Caspers, J. M. Kros, and G. J. Puppels. Tissue characterization using high wave number Raman spectroscopy. *Journal of Biomedical Optics*, 10(3):031116, 2005.

[89] M. G. Shim, L.-M. W. K. Song, N. E. Marcon, and B. C. Wilson. In vivo near-infrared Raman spectroscopy: demonstration of feasibility during clinical gastrointestinal endoscopy. *Photochemistry and Photobiology*, 72(1):146–150, 2000.

[90] U. Utzinger and R. R. Richards-Kortum. Fiber optic probes for biomedical optical spectroscopy. *Journal of Biomedical Optics*, 8(1):121–147, 2003.

[91] H. P. Buschman, G. Deinum, J. T. Motz, M. Fitzmaurice, J. R. Kramer, A. van der Laarse, A. V. Bruschke, and M. S. Feld. Raman microspectroscopy of human coronary atherosclerosis: Biochemical assessment of cellular and extracellular morphologic structures in situ. *Cardiovascular Pathology*, 10(2):69–82, 2001.

[92] A. R. de Paula and S. Sathaiah. Raman spectroscopy for diagnosis of atherosclerosis: a rapid analysis using neural networks. *Medical Engineering and Physics*, 27(3):237–244, 2005.

[93] T. C. Bakker Schut, M. J. H. Witjes, H. J. C. M. Sterenborg, O. C. Speelman, J. L. N. Roodenburg, E. T. Marple, H. A. Bruining, and G. J. Puppels. In vivo detection of dysplastic tissue by Raman spectroscopy. *Analytical Chemistry*, 72(24):6010–6018, 2000.

[94] M. G. Shim, B. C. Wilson, E. Marple, and M. Wach. Study of fiber-optic probes for in vivo medical Raman spectroscopy. *Applied Spectroscopy*, 53(6):619–627, 1999.

[95] Y. Komachi, H. Sato, K. Aizawa, and H. Tashiro. Micro-optical fiber probe for use in an intravascular Raman endoscope. *Applied Optics*, 44(22):4722–4732, 2005.

[96] Y. Komachi, H. Sato, and H. Tashiro. Intravascular Raman spectroscopic catheter for molecular diagnosis of atherosclerotic coronary disease. *Applied Optics*, 45(30):7938–7943, 2006.

[97] Y. Komachi, T. Katagiri, H. Sato, and H. Tashiro. Improvement and analysis of a micro Raman probe. *Applied Optics*, 48(9):1683–1696, 2009.

[98] L. F. Santos, R. Wolthuis, S. Koljenovic, R. M. Almeida, and G. J. Puppels. Fiber-optic probes for in vivo Raman spectroscopy in the high-wavenumber region. *Analytical Chemistry*, 77(20):6747–6752, 2005.

[99] G. J. Tearney and B. E. Bouma. Atherosclerotic plaque characterization by spatial and temporal speckle pattern analysis. *Optics Letters*, 27(7):533–535, 2002.

[100] S. K. Nadkarni, B. E. Bouma, T. Helg, R. Chan, E. Halpern, A. Chau, M. S. Minsky, J. T. Motz, S. L. Houser, and G. J. Tearney. Characterization of atherosclerotic plaques by laser speckle imaging. *Circulation*, 112(6):885–892, 2005.

[101] S. K. Nadkarni, B. E. Bouma, D. Yelin, A. Gulati, and G. J. Tearney. Laser speckle imaging of atherosclerotic plaques through optical fiber bundles. *Journal of Biomedical Optics*, 13(5):054016, 2008.

[102] A. Savitzky and M. J. E. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry*, 36(8):1627–1639, 1964.

[103] C. A. Lieber and A. Mahadevan-Jansen. Automated method for subtraction of fluorescence from biological Raman spectra. *Applied Spectroscopy*, 57(11):1363–7, 2003.

[104] R. Manoharan, J. J. Baraga, M. S. Feld, and R. P. Rava. Quantitative histochemical analysis of human artery using Raman spectroscopy. *Journal of Photochemistry and Photobiology B: Biology*, 16(2):211–33, 1992.

[105] L. Robert and W. Hornebeck. *Elastin and Elastases*, volume II. CRC Press, Boca Raton, FL, 1989.

[106] M. A. Randolph, K. Anseth, and M. J. Yaremchuk. Tissue engineering of cartilage. *Clinics in Plastic Surgery*, 30(4):519, 2003.

[107] P. R. Carey. *Biochemical applications of Raman and resonance Raman specstroscopies.* Academic Press, New York, 1982.

[108] M. Diem. *Introduction to Modern Vibrational Spectroscopy.* John Wiley & Sons, Inc., New York, 1993.

[109] R. Manoharan, Y. Wang, and M. S. Feld. Review: Histochemical analysis of biological tissues using Raman spectroscopy. *Spectrochemica Acta Part A*, 52:215–249, 1996.

[110] F. S. Parker. *Applications of Infrared, Raman, and Resonance Raman Spectroscopy in Biochemistry.* Springer, 1983.

[111] M. R. Prince, G. M. LaMuraglia, and MacNichol, Jr. E. F. Increased preferential absorption in human atherosclerotic plaque with oral beta carotene. Implications for laser endarterectomy. *Circulation*, 78(2):338–344, 1988.

[112] S. W. E. van de Poll, K. Kastelijn, T. C. B. Schut, C. Strijder, G. Pasterkamp, G. J. Puppels, and A. van der Laarse. On-line detection of cholesterol and calcification by catheter based Raman spectroscopy in human atherosclerotic plaque ex vivo. *Heart*, 89(9):1078–1082, 2003.

[113] J. Ma and Y.-S. Li. Fiber Raman background study and its application in setting up optical fiber Raman probes. *Applied Optics*, 35(15):2527–2533, 1996.

[114] A. Roggan, M. Friebel, K. Dorschel, A. Hahn, and G. Muller. Optical properties of circulating human blood in the wavelength range 400–2500 nm. *Journal of Biomedical Optics*, 4(1):36–46, 1999.

[115] S. Waxman, K. Khabbaz, R. Connolly, J. Tang, A. Dabreo, L. Egerhei, F. Ishibashi, J. E. Muller, and G. M. Tearney. Intravascular imaging of atherosclerotic human coronaries in a porcine model: a feasibility study. *International Journal of Cardiovascular Imaging*, 24(1):37–44, 2008.

[116] J. A. Gardecki, J. T. Motz, A. H. Chau, B. E. Bouma, and G. J. Tearney. Establishing laser exposure levels for intracoronary Raman diagnostic procedures. *submitted*, 2009.

[117] S. A. Prahl, M. Keijzer, S. L. Jacques, and A. J. Welch. A Monte Carlo model of light propagation in tissue. *SPIE Proceedings of Dosimetry of Laser Radiation in Medicine and Biology*, IS 5:102–111, 1989.

[118] L. G. Henyey and J. L. Greenstein. Diffuse radiation in the galaxy. *Astrophysical Journal*, 93:70–83, 1941.

[119] A. J. Welch, C. Gardner, R. Richards-Kortum, E. Chan, G. Criswell, J. Pfefer, and S. Warren. Propagation of fluorescent light. *Lasers in Surgery and Medicine*, 21(2):166–178, 1997.

[120] A. M. K. Enejder, T.-W. Koo, J. Oh, M. Hunter, S. Sasic, M. S. Feld, and G. L. Horowitz. Blood analysis by Raman spectroscopy. *Optics Letters*, 27:2004–2006, 2002.

[121] J. Swartling, A. Pifferi, A. M. K. Enejder, and S. Andersson-Engels. Accelerated Monte Carlo models to simulate fluorescence spectra from layered tissues. *Journal of the Optical Society of America A*, 20(4):714–727, 2003.

[122] L.-H. Wang, S. L. Jacques, and L.-Q. Zheng. MCML - Monte Carlo modeling of photon transport in multi-layered tissues. *Computer Methods and Programs in Biomedicine*, 47:131–146, 1995.

[123] J. E. Gentle. *Random Number Generation and Monte Carlo Methods.* Springer, 2003.

[124] B. E. Saleh and M. C. Teich. *Fundamentals of Photonics.* Wiley-Interscience, 2nd edition, 2007.

[125] F. M. Li and A. Nathan. *CCD Image Sensors in Deep-Ultraviolet: Degradation Behavior and Damage Mechanisms.* Springer, 2005.

[126] H. H. Madden. Comments on the Savitzky-Golay convolution method for least-squares-fit smoothing and differentiation of digital data. *Analytical Chemistry*, 50(9):1383–1386, 1978.