12

Quantitative Biological Raman Spectroscopy

Wei-Chuan Shih, Kate L. Bechtel, and Michael S. Feld
George R. Harrison Spectroscopy Laboratory Massachusetts Institute of Technology
Cambridge, MA 02139

12.1 Introduction 366
12.2 Review 367
12.3 Quantitative Considerations for Raman Spectroscopy 370
12.4 Biological Considerations for Raman Spectroscopy 373
12.5 Instrumentation 376
12.6 Data Pre-Processing 378
12.7 *In Vitro* and *In Vivo* Studies 383
12.8 Toward Prospective Application 386
12.9 Conclusion 392
Acknowledgements 392
References 393

Raman spectroscopy is a powerful technique for identifying the molecular composition of materials. It can also be used to quantify the substances present. Recently, quantitative Raman spectroscopy has been used in biological tissue for disease diagnosis and, in blood, to measure concentrations of analytes such as glucose non-invasively. A characteristic feature of biological tissue is its high turbidity, due to the interplay of scattering and absorption. In addition, the complexity of biological tissue results in significant spectral overlap. These factors make the quantification of analyte concentrations difficult. Measurement accuracy can be improved if these difficulties can be overcome. This chapter discusses the application of quantitative Raman spectroscopy to biological tissue. Section 12.1 provides an introduction to Raman spectroscopy. Section 12.2 reviews existing work relevant to quantitative analysis in biological media. Quantitative and biological aspects of Raman spectroscopy are discussed in sections 12.3 and 12.4. Section 12.5 discusses instrumentation, using the instrument developed in our laboratory as an example. Data preprocessing is discussed in section 12.6. In section 12.7 we review our glucose studies in blood serum, whole blood and human subjects. Section 12.8 introduces two new techniques, constrained regularization (CR) and intrinsic Raman spectroscopy (IRS), which are shown to significantly improve measurement accuracy. Additional considerations are discussed in the context of future directions. Section 12.9 concludes the chapter.

Key words: Raman spectroscopy, glucose sensing, turbidity, scattering, absorption, biological tissue, intrinsic Raman spectroscopy (IRS).

12.1 Introduction

12.1.1 Introduction to Raman spectroscopy

Light scattering is a well-known form of the light-matter interaction process. Scattering redirects light incident on an atom or molecule. Most of the scattered light has the same frequency as the incident light, and therefore there is no energy exchange. This process is called elastic scattering and, for scatterers small compared to the wavelength, it gives rise to Rayleigh scattering. A tiny amount of the scattered light, however, is shifted in frequency due to transfer of energy, most commonly vibrational energy, to or from the molecule. The excitation light can set the molecule into vibration at the molecular vibrational frequency, ν_V. This process, called Raman scattering, is an inelastic scattering process, as energy is exchanged between the molecule and the incident light.

From a quantum-mechanical point of view, an incident photon of frequency ν_L, wavelength (c_0/ν_L), and energy $h\nu_L$, with c_0 the speed of light and h Planck's constant, is instantaneously taken up by the molecule, forming a "virtual state" that is usually lower in energy than the electronic transitions of the molecule. A new photon is created and scattered from this virtual state. If the new photon is down-shifted in frequency, the process is called Stokes-Raman scattering [1]. The resulting photon will have a reduced energy $h(\nu_L - \nu_V)$. Similarly, a molecule can begin in an excited vibrational state and proceed, via the virtual state, to the ground state. This generates an up-shifted "anti-Stokes" Raman scattered photon, with an increased energy $h(\nu_L + \nu_V)$. The processes of Rayleigh, Stokes Raman, anti-Stokes Raman with unshifted, down-shifted, and up-shifted frequencies of the scattered light, respectively, are illustrated in Fig. 12.1.

Raman scattering, discovered by Raman and Krishnan in 1928 [2], provides a way to measure molecular composition through inelastic scattering. The frequency shift of the scattered light is a direct measure of the vibrational frequency (*i.e.* energy) of the molecule. Each molecule has its own distinct vibrational frequency or frequencies. The frequency spectrum of the Raman-scattered light thus provides a unique fingerprint of the molecule. The Raman spectrum of material with multiple constituents can thus be used to determine its molecular composition.

A Raman spectrum consists of scattered intensity plotted vs. energy, or frequency, as shown in Fig. 12.2 for glucose in water. Each peak corresponds to a given Raman shift from the incident light energy $h\nu_L$. The energy difference between the

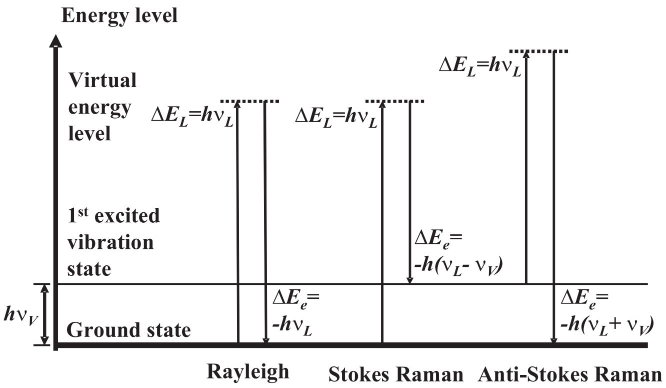

FIGURE 12.1: Energy diagram for Rayleigh, Stokes Raman, and anti-Stokes Raman scattering.

initial and final vibrational states, $h\nu_V$, the Raman shift ν_V, is usually measured in wavenumbers (cm^{-1}), and is calculated as ν_V/c. Raman shifts from a given molecule are always the same, regardless of the excitation frequency (or wavelength). This provides flexibility in selecting a suitable laser excitation wavelength for a specific application.

Infrared (IR) absorption, which probes vibrational structure in the energy range 400–4000 cm^{-1} (25–2.5 μm wavelength range), is also indicative of molecular vibrations. However, these wavelengths are not readily transmitted by most materials. Raman spectroscopy and IR absorption both probe the vibrational structure of molecules, and in many cases the same vibrations are observed. IR absorption is sensitive to vibrational frequencies that change the permanent dipole of the molecule. Raman scattering measures vibrational frequencies that result in a change of polarizability.

Near-infrared (NIR) absorption spectroscopy probes the energy range from 4000 to 10000 cm^{-1} (2.5–1 μm wavelength range), where overtone and combination bands of molecular vibrations occur. Such transitions are quantum mechanically “forbidden,” and are significantly weaker, and with broader features, than those observed in IR absorption. However, in contrast to IR absorption, shorter wavelength NIR light is conveniently transmitted by common optical materials, conferring a substantial advantage over IR absorption in instrumentation. As mentioned earlier, Raman shifts are independent of excitation wavelength, and thus there is flexibility in choosing the wavelength range.

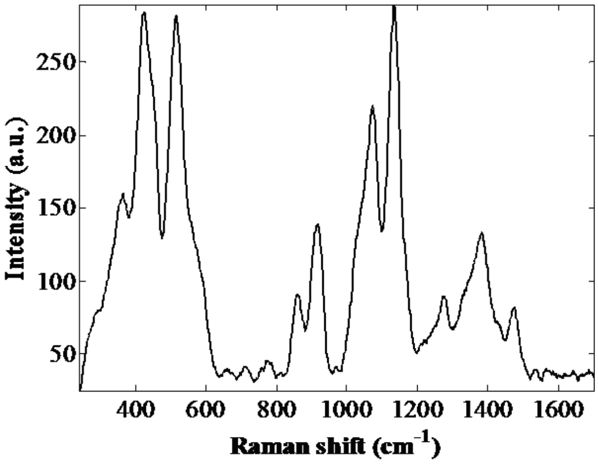

FIGURE 12.2: A Raman spectrum consists of scattered intensity plotted vs. energy. This figure shows an aqueous glucose solution as an example.

12.2 Review

This section briefly reviews the literature relevant to quantitative biological Raman spectroscopy. Raman spectroscopy of biological tissue was initially demonstrated using NIR Fourier transform Raman spectroscopy [3, 4]. In contrast to the visible wavelength range, water absorption and background due to laser-induced autofluorescence are both smaller in the NIR, thus enabling deeper penetration depth and observation of order-of-magnitude weaker Raman peaks. In its early development stages, Raman spectroscopy was primarily employed as a qualitative tool for chemical identification, with limited ability for quantification. Through the introduction and improvement of lasers, CCDs and other optical components, quantitative analysis became possible.

In the following, we review three categories of work: semi-quantitative, univariate, and multivariate analyses. Such distinctions are made based on the types of analysis carried out. For instance, the hallmark of semi-quantitative work is normalization to the overall peak intensity. Although absolute intensity information is lost in the normalization step, quantitative analysis can be applied afterwards. Univariate analysis uses one or a few characteristic peaks through measurement of peak heights or integration of the area under the peaks.

In contrast to univariate analysis, multivariate techniques are often called “full spectral range” methods. This type of analysis is usually carried out when spectral overlap exists, and therefore the characteristic peaks of interest are not obvious or are "contaminated" by adjacent features not belonging to the substance of interest. As the analysis becomes more sophisticated and more multivariate in nature, the issue of model robustness must be considered. We discuss various aspects of this topic in the subsection 12.3.1.

12.2.1 Semi-quantitative implementation

Raman spectroscopy has been employed in disease diagnosis using morphological models, with the rationale that characteristic morphological features are representative disease biomarkers. This approach is based on the unique correspondence between a particular morphological structure and the underlying chemical substance. These models are constructed via ordinary least squares (OLS) analysis, which assumes that the (important) spectral components are all precisely known, and that the observed experimental spectrum can be represented as a linear superposition of these spectral components, weighted by their concentrations [5]. Haka *et al.* [6] developed a morphological model for breast cancer diagnosis using confocal Raman microscopy. They analyzed the Raman spectral features of normal, benign and malignant tissue samples in terms of the relative amount of collagen, fat, keratin, etc. Van de Poll *et al.* [7] and Buschman *et al.* [8] studied atherosclerosis using a similar approach. Spectra of individual morphological structures were obtained using confocal Raman microscopy, and then applied to fit tissue spectra collected with an optical fiber probe. In these studies, spectra of both the model components and those taken in tissue were normalized to their respective highest peaks, and absolute intensity information was not retained. However, the chemical composition could be quantified in terms of the relative proportions of the model components and then correlated with disease. For example, the relative quantity of collagen to fat was found to be a relevant breast cancer biomarker.

12.2.2 Univariate implementation

For some applications, characteristic and distinct Raman peaks of the chemical (molecule) of interest can be observed with little difficulty. Peak height measurement or integration of the area under the peak can be used as a quantitative indicator of the substance. Caspers *et al.* [9] developed confocal Raman microscopy to perform non-invasive determination of the water profile in human skin *in vivo*. Peterson *et al.* [10] reported the acquisition of whole blood Raman spectra *in vivo* using tissue modulation. Glucose concentrations were subsequently extracted from the area under particular spectral peaks of the whole blood spectra. A calibration model derived from one individual was then used to generate meaningful predictions on independent data.

12.2.3 Multivariate implementation

In the Raman spectra of more complicated chemical systems, the various underlying components (called "interferents") generally exhibit substantial spectral overlap.

Therefore, no distinct Raman peak is available for peak height or area measurement, and the full spectrum must be used. This is called multivariate analysis. The goal of multivariate analysis is to obtain a spectrum of numbers, $\mathbf{b}(j)$, with j the wavelength index. When $\mathbf{b}(j)$ is projected onto an experimental spectrum $\mathbf{s}(j)$, one obtains accurate prediction of the analyte's concentration, c [5]. Such spectra are often described as column vectors, with each dimension corresponding to a given sampling point on the wavelength axis. In these terms, c is obtained as the scalar product of **b** with the experimental spectrum, **b**:

$$c = \mathbf{s}^T \mathbf{b} \tag{12.1}$$

where lowercase boldface type denotes a column vector, and the superscript T denotes the transpose. Note that Eq. (12.1) assumes linearity, *i.e.*, the observed spectrum can be represented as a linear superposition of underlying spectral components. Multivariate analysis proceeds in two steps. In calibration, one correlates known concentrations with spectra to obtain **b**. The resulting **b**, sometimes called the regression vector, is then used to predict the concentration of an unknown sample. Multivariate calibration is further discussed in detail in subsection 12.3.1, below.

Non-invasive measurement of blood analyte concentrations is a widely pursued topic, and most studies employ multivariate techniques to extract analyte-specific concentration information. However, from a data analysis standpoint multivariate calibration presents more challenges than univariate methods, because of system complexity and the resulting spectral overlap. Owing to its potential impact on diabetes, glucose has been often used as a model analyte.

In vitro measurements of glucose have been performed in filtered blood serum [11, 12], blood serum [13], and whole blood [14]. Rohleder *et al.* [12] discovered that measurements from serum are greatly improved by ultrafiltration to remove macromolecules that cause intense Raman background and subsequently impair measurement accuracy. Results from whole blood were found to have greater error than those from filtered or unfiltered serum, but were still within the clinically acceptable range. Lambert *et al.* [15] measured human aqueous humor, simulating measurements in the eye, a convenient target for optical techniques. Our group studied glucose non-invasively in human subjects using Raman spectroscopy coupled with multivariate analysis; Enejder *et al.* [16] accurately measured glucose concentrations in 17 non-diabetic volunteers following an oral glucose tolerance protocol. Results based on analysis of spectra from individual and multiple volunteers indicated that the calibration model was based on glucose rather than spurious correlations.

12.3 Quantitative Considerations for Raman Spectroscopy

Traditionally, Raman spectroscopy has been utilized as an analytical tool for chemical identification and fingerprinting, where analysis has been based on observation of characteristic Raman peaks. As mentioned in the previous section, many more challenges are encountered when Raman spectroscopy is used as a quantitative analytical tool. We discuss the challenges below.

12.3.1 Considerations for multivariate calibration models

As discussed previously, although Raman spectroscopy provides good molecular specificity, spectral overlap is inevitable with the presence of multiple constituents. Since the glucose Raman signal is only $\sim$0.3% of the total skin Raman signal [17, 18], and the spectrum is complicated by shot noise and varying fluorescence background, multivariate calibration is necessary. There are two types of multivariate techniques, explicit and implicit. In explicit techniques such as OLS [5], the $\mathbf{b}$ vector is calculated from the full set of known spectral components. In implicit techniques, such as partial least squares (PLS) [19, 20] analysis, the $\mathbf{b}$ vector is derived from a calibration data set composed of samples with known concentrations of the analyte of interest.

Since multivariate calibration models are often built on an underdetermined data set, careful assessment of model validity is required. Here we present some considerations for evaluating a calibration model. The reader is referred to the references for more detailed information about multivariate calibration.

12.3.2 Fundamental and practical limits

In spectroscopy, the amplitude of the Raman spectrum of the analyte of interest depends on the number of analyte molecules sampled by the incoming light. The effective path length (in transmission mode) and sampling volume (in reflection mode) of the light are important parameters in estimating detection limits in turbid media. Modeling techniques such as diffusion theory [21] and Monte Carlo simulation [22] can be employed to calculate the fluence distribution inside the sample and the angular and radial profiles of the transmitted or reflected flux. Simulations with synthetic data or experiments employing tissue-simulating physical models (called “phantoms”) can be of great value in determining how close the theoretical limit can be realized in practice. In these studies, experimental conditions (*e.g.*, signal-to-noise ratio (SNR), instrumental drifts) and tissue phantom composition (*e.g.*, interferents, concentrations) can be precisely controlled and well characterized in advance. Demonstrating that the chosen technique and instrument can measure physiological levels of the analyte of interest in phantoms is necessary but not sufficient to validate *in vivo* results. *In vivo* calibration models can only be validated by prospective studies.

12.3.3 Chance or spurious correlation

Multivariate calibration algorithms are powerful, yet can be misleading if used without caution. Owing to the nature of the underdetermined data set, minute correlations present in the data set can be misinterpreted by the algorithm as actual analyte-specific variations. For example, Arnold *et al*. [23] measured the NIR absorption spectra of tissue phantoms devoid of glucose, and used temporal glucose concentration profiles published by different research groups to demonstrate that the calibration model could produce an apparent correlation with glucose even though none was present. It is important to note that calibration results such as these satisfied multiple criteria for judging the validity of a calibration model.

Overfitting is another cause for spurious correlations. In multivariate calibration, a large number of sample spectra can be reduced to fewer factors. In practice, only a subset of factors is significant in modeling the underlying analyte variations, while others are more likely to be dominated by noise and measurement errors. Although an apparently lower calibration error may be obtained by including more factors in the calibration model, the reduction in error may be fortuitous and the resulting model may have less predictive capability.

The lesson here is that chance or spurious correlations may be inadvertently incorporated in the calibration model even when rigorous validation procedures have been followed. Additionally, if these chance or spurious correlations exist in prospective data, even good prediction results can be based on non-analyte-specific effects. Incorporating prior or additional information into the calibration model has been shown to provide more immunity to chance correlations [24–27].

12.3.4 Spectral evidence of the analyte of interest

The difficulty in visualizing analyte-specific information in biological spectra makes it challenging to verify the origin of the spectral information used by the calibration model and confirm that positive results are actually based on the analyte of interest. However, some of this information can be obtained by examining the **b** vector. The **b** vectors obtained from spectroscopic data contain spectral information about all the components of the model. Under ideal, noise-free conditions, **b** can be explicitly derived from the model component spectra (via OLS), or implicitly obtained from the calibration data set. This ideal **b** vector is also termed the net analyte signal [28, 29]. As mentioned above, **b** is identical (within a scale factor) to the pure component spectrum of the analyte of interest if no other interferents are present. In other words, **b** should "look" progressively more like the glucose spectrum as model complexity decreases. For example, when spectral overlap is low, as in Raman spectroscopy, spectral features of glucose can be identified in the experimentally derived **b** vector as supporting evidence that the model is based on glucose rather than spurious correlations [13, 14, 16].

Although a complete model is virtually never available for *in vivo* experiments, a good approximation to the actual **b** vector is often obtainable from an OLS estimate. Therefore, it is always useful to compute a theoretically "correct" **b** vector and compare it to the experimentally derived **b**. If the two significantly differ, the discrepancy should be investigated.

12.3.5 Minimum detection limit

If all component spectra in a mixture sample are known, the minimum detection error, Δc, can be calculated via a simple formula derived by Scepanovic *et al.* [30] of our laboratory:

$$\Delta c = \frac{\sigma}{\|\mathbf{s}_k\|}\mathrm{OLF}_k, \tag{12.2}$$

where the subscript k designates the constituent analyte to be measured. The first factor on the right hand side, σ, is the standard deviation of the noise in the measured spectrum. The second factor, $\|\mathbf{s}_k\|$, is the signal strength of the analyte, represented as the norm of its Raman spectrum. The last factor, OLF_k, is termed the "overlap factor," and can take on values between 1 and infinity. OLF_k indicates the amount of overlap between the Raman spectrum of the analyte and those of the interferents. Mathematically, OLF_k is the inverse of the correlation coefficient between the analyte spectrum and the OLS regression vector ($\mathbf{b}_{\mathrm{OLS}}$):

$$\mathrm{OLF}_k = \frac{1}{\mathrm{corr}(\mathbf{b}_{\mathrm{OLS}}, \mathbf{s}_k)}. \tag{12.3}$$

$\mathbf{b}_{\mathrm{OLS}}$ is the portion of the analyte spectrum $\mathbf{s}_k$ that is orthogonal to all interferents. When there are no interferents, $\mathbf{b}_{\mathrm{OLS}}$ is identical to the analyte spectrum and thus $\mathrm{OLF}_k = 1$. When interferents are present, $\mathrm{corr}(\mathbf{b}_{\mathrm{OLS}}, \mathbf{s}_k)$ is always smaller than one and therefore OLF_k is always larger than 1. To estimate the overlap factor for glucose measurements in skin, we built a 10-constituent model, approximating the Raman spectrum of human skin. Starting with only glucose, the Raman spectra of other constituents, including, collagen type I, keratin, triolein, actin, collagen type III, cholesterol, phosphatidylcholine, hemoglobin and water were progressively added, increasing model complexity. The correlation between $\mathbf{b}_{\mathrm{OLS}}$ and the glucose spectrum was reduced from 1 to 0.73, as shown in Fig. 12.3, indicating that the detection limit indeed becomes worse for more complex chemical systems.

Equation (12.2) provides a practical way to estimate the minimum detection limit, Δc, based on easily obtainable experimental parameters. Lower fluorescence background introduces less shot noise and therefore increases the Raman SNR. High molecular specificity in the Raman spectra leads to less spectral overlap and thus reduces the OLF for a specific analyte. A practical estimate of the minimum detectable glucose concentration in blood for our instrument (section 12.5) is $\Delta c \sim 2$ mg/dL.

12.4 Biological Considerations for Raman Spectroscopy

12.4.1 Using near infrared radiation

Raman shifts are independent of excitation wavelength and thus offer flexibility in choice of wavelength range. NIR excitation provides three advantageous features for

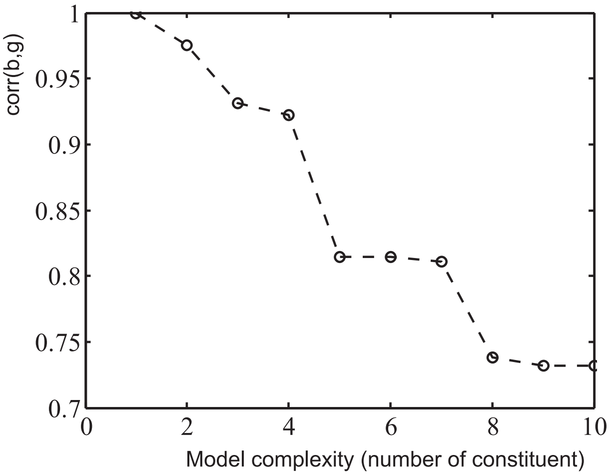

FIGURE 12.3: Correlation between the OLS regression vector ($\mathbf{b}_{\mathrm{OLS}}$) and the glucose spectrum versus model complexity.

studying biological tissue: low-energy optical radiation, deep penetration depth, and reduced background fluorescence. Excitation wavelengths in the NIR region prevent hazardous ionization of tissue constituents. The lack of prominent absorbers in the NIR region enables sampling over sufficient depth, of the order of ~ 1 mm. The reduced shot noise associated with the low fluorescence background induced by NIR excitation provides an order of magnitude improvement in sensitivity in extracting Raman signals. As a result, our Raman studies of biological tissue employ 830 nm as the excitation wavelength. As shown in Fig. 12.4, this provides Raman spectra over 830-1000 nm, a wavelength range in the "diagnostic window" (deep penetration), in which silicon-based charge coupled device (CCD) detectors have very high quantum efficiency.

Figure 12.4 illustrates the absorption spectra of major endogenous tissue absorbers, water, skin melanin, hemoglobin, and fat. Also shown is the scattering spectrum of 10% Intralipid, a lipid emulsion often used to simulate tissue scattering. The diagnostic window is indicated by the rectangle.

12.4.2 Background signal in biological Raman spectra

Raman spectra of biological samples are often accompanied by strong background. The source of the background is often attributed to fluorescence, particularly when UV/visible laser excitation is employed. Proteins and lipids present in biological tissue can contribute to the fluorescence background [11]. The autofluorescence of skin from endogenous fluorophores with UV/visible light excitation is well known and, in fact, it has been applied to the diagnosis of disease states such as psoriasis [31]

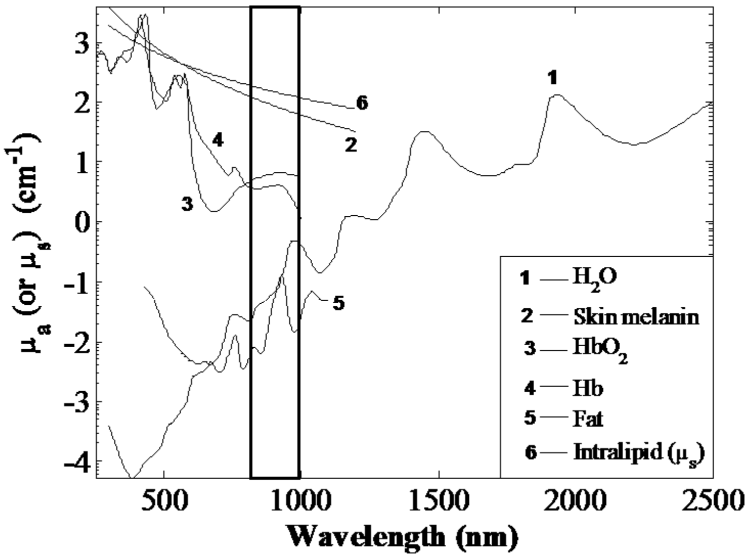

FIGURE 12.4: Absorption spectra of water, skin melanin, hemoglobin, and fat. Also shown is the scattering spectrum of 10% Intralipid, a lipid emulsion often used to simulate tissue scattering. Data are obtained from http://omlc.ogi.edu/spectra/index.html.

and diabetes. In these cases, collagen fluorescence changes owing to glycation, the process in which a single sugar such as glucose binds to a protein or lipid molecule without the controlling action of an enzyme [32]. Furthermore, our laboratory utilizes the autofluorescence of epithelial tissue components to diagnose dysplasia [33]. The shot noise accompanying this fluorescence background limits the detection capability.

Time-dependent background variations influence subsequent multivariate analysis. In a human skin study using UV/visible light excitation, Zeng *et al.* [34] described the temporal background intensity decay as fluorescence photobleaching. They fit the intensity decay to a double-exponential function, with the time constants ascribed to photobleaching rates of different fluorophores in the stratum corneum and dermis. On the other hand, Jongen and Sterenborg [35] found that even for a single fluorophore, the turbidity of tissue causes fluorescence decay to deviate from single-exponential behavior. Hence, double-exponential behavior need not be ascribed to two fluorophores. The reasoning for this argument is that the fluorescence signal from a multi-layered turbid medium is the sum of the contributions from each layer. Fluorescence from a deeper lying layer appears weaker and will photobleach at a slower rate because of diminished laser power. Thus, the relative contribution of fluorescence from deeper layers will appear as smaller signals that decay slower, whereas the superficial layers will exhibit stronger signals that decay faster. This illustrates the strong influence that the optical properties of the sample can have on the observed behavior of light.

While implicit multivariate calibration techniques can remove the detrimental effects of the background to some extent, their efficacy is impaired. Thus, it is desirable to either reduce the background during data collection or subtract it via modeling, without introducing artifacts. Most background subtraction methods in the literature are based on polynomial fitting. Since the background has little structure, a slowly-varying low-order polynomial can characterize it [4, 36–38]. Refs. [36–38] find that a fifth-order polynomial is the most effective fit to the background.

12.4.3 Heterogeneities in human skin

Uniform analyte distribution is often a good assumption for liquid samples such as serum or whole blood with continuous stirring. However, for biological tissue, human skin in particular, heterogeneity is a major feature. Detailed morphological structure and molecular constituents associated with skin heterogeneity have been studied using confocal Raman spectroscopy [39]. Skin has two principal layers: epidermis and dermis. The epidermis is the outmost layer, and itself consists of multiple layers including the stratum corneum as the major component, stratum lucidum, and stratum granulosum. The major chemical constituent of human epidermis is keratin, comprising approximately 65% of the stratum corneum. The dermis is also a layered structure composed mainly of collagen and elastin. Blood capillaries are present in the dermis, and thus this region is targeted for optical analysis for glucose detection. However, it has been suggested that the majority of the glucose molecules sampled by optical techniques arise from the interstitial fluid (ISF), which is primarily found at the epidermis-dermis interface [40]. The role of skin heterogeneity in non-invasive measurement of blood analyte concentrations is an important factor that has not yet been fully studied.

12.5 Instrumentation

As discussed previously, background fluorescence from biological tissue impedes observation of Raman signals with UV/visible excitation wavelengths. To overcome this limitation, the use of NIR excitation was introduced, initially using Fourier-transform spectrometers [3]. With the advent of high quantum efficiency CCD detectors and holographic diffractive optical elements, our group and others have led the way in switching to CCD-based dispersive spectrometers [4, 10, 11, 13-15]. The advantages of dispersive NIR Raman spectroscopy are that compact solid-state diode lasers can be used for excitation, the imaging spectrograph can be f-number matched with optical fibers for greater throughput, and cooled CCD detectors offer shot-noise-limited detection. As a tutorial for the selection of building blocks for a Raman instrument with high collection efficiency, we present a summary of the key design considerations.

12.5.1 Excitation light source

Laser excitation at one of two wavelengths, 785 and 830 nm, is most common. The tradeoff is that excitation at lower wavelengths has a higher efficiency of generating Raman scattering but also generates more intense background fluorescence. The current trend is towards the use of external cavity laser diodes because they are compact and of relatively low cost. Other researchers use argon-ion laser pumped titanium-sapphire lasers. The titanium-sapphire laser provides higher power output with broader wavelength tunability, but it is bulkier and much more expensive than diode lasers.

Narrow-band excitation must be used to prevent broadening of the Raman bands. Further, the wings of the laser emission (amplified spontaneous emission) can extend beyond the cutoff wavelength of the notch filter used to suppress the elastically scattered light, and can obscure low wavenumber Raman bands. This problem is most severe with high power diode lasers, and a holographic bandpass or interference laser line filter with attenuation greater than 6OD is usually required. Lastly, for quantitative measurements a photodiode is often needed to monitor the laser intensity to correct for power variations.

12.5.2 Light delivery, collection, and transport

The filtered laser light can be delivered to the sample either via free-space transmission or through an optical fiber. In free-space embodiments, beam shaping is usually performed to correct for astigmatism and other laser light artifacts. The incident light at the sample can be either focused or collimated, depending on collection considerations. For biological tissue, the total power per unit area delivered to the tissue must be limited for safety, and thus spot size on the tissue is an oft-reported parameter.

Raman probes constructed from fused silica optical fibers have gained much attention recently. Typically, low-OH content fibers are utilized to reduce the fiber fluorescence. The probe design also includes filters at the distal end to suppress the fused silica Raman signal from the excitation fiber and prevent the elastically scattered light entering the collection fibers [41]. Commercial probes are now available, and they offer ruggedness and easy access to samples with geometrical or other constraints.

As Raman scattering is a weak process, high collection efficiency is required. Specialized optics such as Cassegrain microscope objectives and non-imaging paraboloidal mirrors have been employed to increase both the collection spot diameter and the effective numerical aperture of the optical system [40]. The majority of light that exits the air-sample interface is elastically scattered at the laser wavelength. This light must be properly attenuated or it will saturate the CCD detector. Holographic or interference notch filters are extensively employed for this purpose, and can attenuate the elastically scattered light to greater than 6OD, while passing the Raman light with greater than 90% efficiency. However, notch filters are very sensitive to the incident angle of the light, and thus provide less attenuation for off-axis light. In some instances the diameter of the notch filter is one of the determining factors of the throughput of an instrument.

Specular reflection, light that is elastically scattered without penetrating the tissue, is also undesirable. Strategies such as oblique incidence [42], 90-degree collection geometry [11], and bringing in excitation light via a hole in the collection mirror have all been employed to reduce its effect [27]. After filtering out most of the elastically scattered light, the Raman scattered light must be transported to the spectrograph with minimum loss. To match the shape of the entrance slit of a spectrograph, the circular shape of the collected light can be relayed by an optical fiber bundle with the receiving end arranged in a round shape and the exiting end arranged in a line [14].

12.5.3 Spectrograph and detector

In dispersive spectrographs for Raman spectroscopy, transmission holographic gratings are often used because of their compactness and high dispersion. Holographic gratings can be custom-blazed for specific excitation wavelengths and provide acceptable efficiency. In addition, liquid nitrogen cooled and, more recently, thermoelectric cooled CCD detectors offer high sensitivity and shot-noise-limited detection in the near infrared wavelength range up to $\sim 1\ \mu$m. These detectors can be controlled using programs such as Labview to facilitate experimental studies.

To increase light throughput in Raman systems, the CCD chip size can be increased vertically to match the spectrograph slit height. However, large format CCD detectors show pronounced slit image curvature that must be corrected for in preprocessing (described in subsection 12.6.1).

As an example of these design considerations, Fig. 12.5 shows a schematic of the high-throughput Raman instrument currently used in our laboratory. We opted for free space delivery of the excitation light through a small hole in an off-axis half-paraboloidal mirror. Backscattered Raman light is collimated by the mirror and passed through a 2.5-inch-diameter holographic notch filter to reduce elastically scattered light. The Raman light is focused onto a shape-transforming fiber bundle, with the exit end serving as the entrance slit of an f/1.4 spectrometer. The pre-filtering stage of the spectrometer was removed to reduce fluorescence and losses from multiple optical elements. The back-thinned, deep depletion, liquid nitrogen-cooled CCD, 1300×1340 pixels, is height-matched to the fiber bundle slit. This instrument was specifically designed for high sensitivity measurements in turbid media.

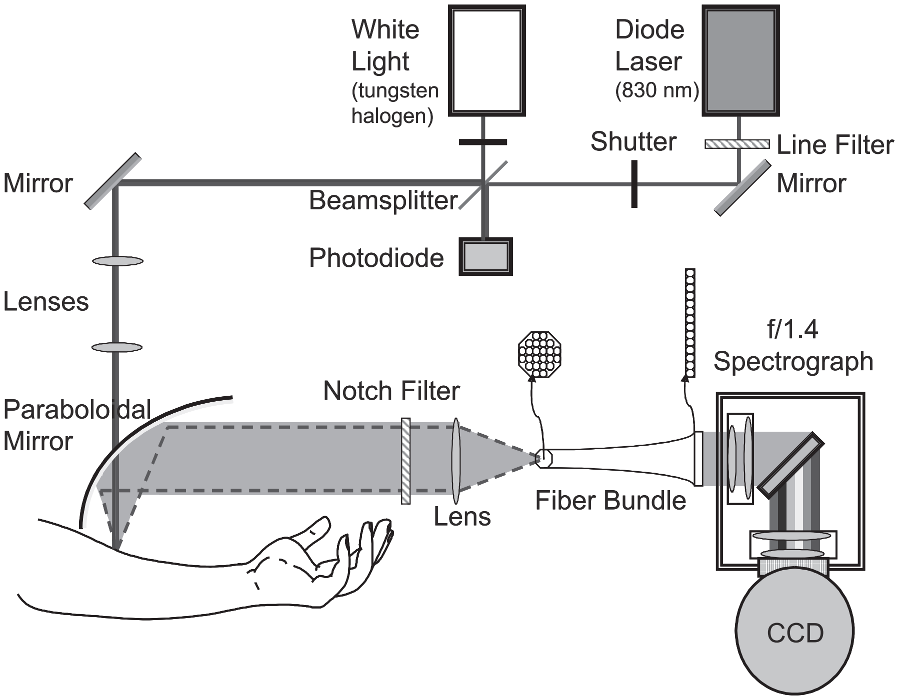

FIGURE 12.5: Schematic of a free space Raman instrument for non-invasive glucose measurements used at the MIT Spectroscopy Laboratory. The white light source is for diffuse reflectance measurements and is not always present.

12.6 Data Pre-Processing

After data collection, various pre-processing steps are undertaken to improve data quality. The pre-processing steps chosen can lead to different calibration results. Therefore, it is important to carefully consider the exact procedures used. The major pre-processing steps are described in the following.

12.6.1 Image curvature correction

Increasing the usable detector area is an effective way to improve light throughput in Raman spectroscopy employing a multi-channel dispersive spectrograph. However, owing to out-of-plane diffraction, the entrance slit image appears curved [43]. Direct vertical binning of detector pixels without correcting the curvature results in degraded spectral resolution.

Various hardware approaches, such as employing curved slits [42, 43] or convex spherical gratings, have been implemented. In the curved slit approach, fiber bundles are employed as shape transformers to increase Raman light collection efficiency. At the entrance end the fibers are arranged in a round shape to accommodate the focal spot, and at the exit end in a curved line, with curvature opposite to that introduced by the remaining optical system.This exit arrangement serves as the entrance slit of the spectrograph and provides immediate first order correction of the curved image, as described below. However, for quantitative Raman spectroscopy, with substantial change of the image curvature across the wavelength range of interest ($\sim$ 150 nm) and narrow spectral features, this first order correction is not always satisfactory.

As an alternative to the hardware approach, software can be employed to correct the curved image, with potentially better accuracy and flexibility for system modification. In past research, we developed a software method using a highly Raman-active reference material to provide a sharp image on the CCD [44]. Using the curvature of the slit image at the center wavelength as a guide, we determine by how many pixels in the horizontal direction each off-center CCD row needs to be shifted in order to generate a linear vertical image. This pixel shift method, as well as the curved-fiber-bundle hardware approach, ignores the fact that the slit image curvature is wavelength dependent. The resulting spectral quality of the pixel shift method is thus equivalent to the curved-fiber-bundle hardware approach [43]. This issue becomes more important when large CCD chips and high-NA spectrographs are employed for increasing the throughput of the Raman scattered light.

Recently, a software approach using multiple polystyrene absorption bands was developed for infrared spectroscopy [16]. In this section we present a similar method that we developed concurrently, which calibrates on multiple Raman peaks to generate a curvature map.

This curvature mapping method shows significant improvement over first-order correction schemes. The curvature mapping method requires an initial calibration step. In this step, a full-frame image is taken with a reference material that has prominent peaks across the spectral range of interest, for example, acetaminophen (Tylenol) powder. We chose nine prominent peaks across the wavelength range of interest, indicated by the arrows in Fig. 12.6.

The calibration algorithm generates a map of the amount of shift for each CCD pixel and a scale factor to maintain signal conservation in each CCD row. Once the map and the scale factor are generated, usually when the system is first set up, the correction algorithm can be applied to future measurements.

Figure 12.7 shows that significant improvement is obtained from the pixel shift method to the curvature mapping method, especially toward either side of the CCD (compare Figs. 7(c) and 7(e)). The overall linewidth reduction in 14 prominent peaks is 7% (FWHM). Such improvement is significant, considering that the equivalent slit width is $\sim$360 μm. If a narrower slit is employed for better spectral resolution, the overall linewidth reduction is even more significant. Note that the images were taken with 5-pixel CCD hardware vertical binning to reduce the amount of data, since the curvature is barely noticeable within such a short range. The error introduced by the hardware binning is much less than 1 pixel, and thus negligible.

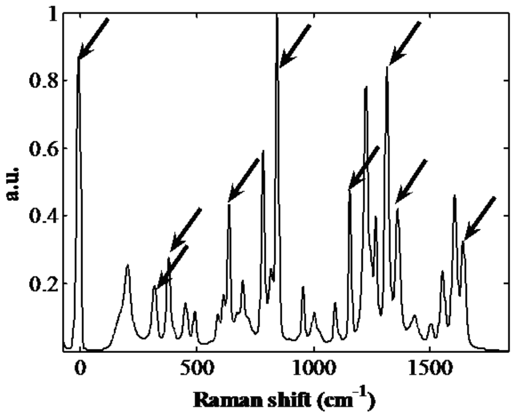

FIGURE 12.6: Raman spectrum of acetaminophen powder, used as the reference material in the calibration step. Nine prominent peaks used as separation boundaries are indicated by arrows.

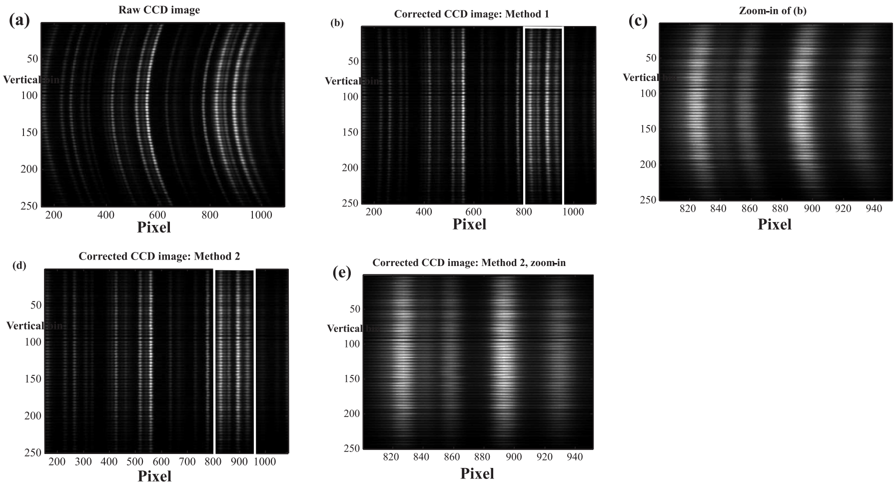

FIGURE 12.7: CCD image of acetaminophen powder. Images were created with 5-pixel hardware binning. (a) Raw image; (b) after applying pixel shift method; (c) zoom-in of the box in (b); (d) after applying curvature mapping method; (e) zoom-in of the box in (d).

12.6.2 Spectral range selection

Multivariate calibration methods attempt to find spectral components based on variance. The presence of a spectral region with large non-analyte-specific variations may bias the algorithm and cause smaller analyte-specific variance to be overlooked. Therefore, the 'fingerprint' region from approximately 300-1700 cm^{-1} is often chosen for analysis.

12.6.3 Cosmic ray removal

Cosmic rays traverse the CCD array at random times with arbitrary intensities, resulting in spikes at individual pixels. When the array is summed and processed, sharp spectral features of arbitrary intensities may appear in the Raman spectra. These artifacts are typically removed before multivariate calibration.

One approach is based on the assumption that the spectrum does not change its intensity from frame to frame other than due to noise and cosmic rays. Therefore, by comparing multiple neighboring frames, a statistical algorithm can be used to identify cosmic rays. Another solution compares adjacent pixels in the same spectrum and detects abrupt jumps in intensity from pixel to pixel. Once a cosmic ray contaminated pixel is identified, its value can be replaced by the average of neighboring pixels.

12.6.4 Background subtraction

As mentioned in section 12.4, the background signal in Raman spectra is a limiting factor in determining the detection limit. Background subtraction techniques only approximate the shape of the background, and thus place a limit on extracting information. The contrasting approach is to not subtract that background, and instead rely on multivariate calibration algorithms to process both the Raman signals and the background. However, for qualitative analysis, background-subtracted spectra provide better interpretation of the underlying constituents.

12.6.5 Random noise rejection and suppression

Photon shot-noise-limited performance can be achieved using a liquid nitrogen cooled CCD camera. When a detector is shot noise limited, the noise can be estimated as the square root of the number of photoelectrons collected in the integration time window. The most effective way to increase the SNR under shot-noise-limited conditions is to increase the integration time of the CCD or the throughput of the instrument. However, extending the integration beyond a certain time scale offers less additional benefit, as other errors begin to dominate performance [40]. Once the data are collected, signal processing is the only way to further enhance the SNR. Pixel binning along the wavelength axis is one method of increasing the SNR, and results indicate that there is an optimal number of binned pixels for optimizing SNR [40]. However, the drawback to this method is degradation in spectral resolution. More commonly employed are Savitzky-Golay smoothing algorithms, which also improve the SNR while better retains spectral resolution.

12.6.6 White light correction and wavelength calibration

When spectra collected from different instruments or on different days are to be compared, white light correction and wavelength calibration are required. White light correction is performed by dividing the Raman spectra by the spectrum from a calibrated light source, for example a calibrated tungsten-halogen lamp measured under identical conditions. The combined spectral response of the optical components, the diffraction grating, and the CCD camera can be effectively removed, thus revealing more of the underlying Raman spectral features. Wavelength calibration is performed to transform the pixel-based axis into a wavelength-based (or wavenumber-based) axis, enabling comparison of Raman features across instruments and time.

12.6.7 Wavelength selection

Although in most experiments Raman spectra are acquired over a continuous wavelength range, analyte-specific information can be distributed non-uniformly across this range. In addition, the overlap factor can change if different wavelengths are chosen for multivariate calibration. Further, because the background is usually non-uniform in wavelength, the shot noise is usually not constant across the entire spectral range. These factors, considered in combination, suggest that there may be advantages when particular wavelength channels (*e.g.*, CCD pixels) are excluded from the spectra. The theoretical basis of wavelength selection and algorithms to perform such selection have been studied [45]. In our laboratory, wavelength selection has not been implemented, but we will consider it in future studies.

12.7 *In Vitro* and *In Vivo* Studies

In this section, we review the application of quantitative Raman spectroscopy to measure blood analytes, *e.g.*, in particular the glucose studies performed in our laboratory. *In vitro* studies have been performed using unprocessed human blood serum and human whole blood. *In vivo* studies have been performed with human volunteers under glucose tolerance test protocols.

12.7.1 Model validation protocol and summary statistics

In our studies, leave-one-out cross validation has been employed as a way to use small data sets efficiently. Validation of the calibration model is crucial before prospective application. Two types of validation schemes can be employed, internal and external. Internal validation, or cross validation, is used when the number of calibration samples is limited. In cross validation, a small subset of calibration data is withheld in the model building step, from which the **b** vector is obtained. The model is then applied to the withheld data. A different subset of calibration data is next withheld and the model rebuilt, and again applied to the withheld data. This process is repeated until all the combinations have been studied.

In determining the optimal model via cross validation, the root mean square error of cross validation (RMSECV) is calculated. RMSECV is defined as the square root of the sum of the squares of the differences between extracted and reference concentrations. The RMSECV is calculated for a particular choice of the number of model parameters, *e.g.*, the number of factors in PLS. An iterative algorithm is often employed to vary the number of parameters and recalculate the RMSECV. The statistically significant minimum RMSECV and the corresponding number of model parameters are then chosen for determining the optimal calibration model.

Various strategies can be employed for grouping spectra for calibration and validation. For example, a single sample can be withheld in a “leave-one-out” scheme, and the calibration and validation process repeated as many times as the number of samples in the calibration data set. In general, “leave-n-out” cross validation can be implemented with n random samples chosen from a pool of calibration data.

When the calibration data set is sufficiently large, external validation, *i.e.*, predictive testing, can be employed. As opposed to internal validation, external validation tests the calibration model and optimizes the number of model parameters on data that never influences the model, and therefore provides a more objective measure than internal validation.

The **b** vector obtained by the validation procedure can be employed prospectively to predict concentrations of the analyte of interest in independent data. Similar to the calculation of RMSECV, root mean square error of prediction (RMSEP) for an independent data set is defined as the square root of the sum of the squares of the differences between predicted and reference concentrations. For feasibility studies, RMSECV is a good indicator of performance as long as the number of calibration samples is statistically sufficient. RMSEP, on the other hand, provides the objective metric by which a technology ultimately must be evaluated.

12.7.2 Blood serum

Our laboratory began investigating non-invasive blood analysis using Raman spectroscopy in the mid 1990’s [46–48]. The first biological sample study was conducted on serum and whole blood samples from 69 patients over a seven-week period [13]. No sample processing or selection criteria were employed, with the exception of locating a few samples with extreme glucose concentrations to represent the range of diabetes patients’ glucose levels. An 830 nm diode laser was employed for excitation and a microscope objective for light collection. The laser power at the sample was $\sim$ 250 mW, and the integration time for each spectrum was equivalent to 300 seconds. The glucose measurement results in serum were quite encouraging, with PLS calibration providing an RMSECV of 1.5 mM. However, the glucose measurement results in whole blood result were not satisfactory because of reduced signals from the highly turbid samples. Glucose spectral features were identified in both the PLS weighting vector and the **b** vector, supporting that the calibration model was based on glucose.

12.7.3 Whole blood

The main difficulty in measuring analyte concentrations in whole blood as compared to serum was attributed to the much higher absorption and scattering of whole blood because of the presence of hemoglobin and red blood cells, respectively. The combined effect resulted in a factor of four decrease in collected Raman signal size.

A subsequent whole blood study in our laboratory by Enejder *et al.* [14] confirmed this hypothesis. A four-fold increase in Raman signal collection size was achieved by employing a paraboloidal mirror and a shape-transforming fiber bundle for better collection efficiency, as depicted in Fig. 12.5. Accurate measurement of multiple analytes was then demonstrated in 31 whole blood samples with laser intensity and integration time similar to the previous serum study [13]. PLS leave-one-out cross validation was performed, and an RMSECV of 1.2 mM was obtained. The ratio of the number of PLS factors to that of the samples ($\sim$1:3) raised the concern of overfitting. However, glucose spectral features were identified in the regression vector, providing confidence that the model was based on glucose.

12.7.4 Human study

Enejder *et al.* [16] conducted a transcutaneous study on 17 non-diabetic volunteers using a version of the instrument depicted in Fig. 12.5 without the white light source. The objective of this study was to provide an initial evaluation of the ability of Raman spectroscopy to measure glucose non-invasively, with the focus on determining its capability in a range of subjects. Spectra were collected from the forearms of human volunteers in conjunction with an oral glucose tolerance test protocol, involving overnight fasting followed by the intake of a high-glucose containing fluid, after which the glucose levels are elevated to more than twice that found under fasting conditions. Periodic reference glucose concentrations were obtained from finger-stick blood samples and subsequently analyzed by a Hemocue portable glucose meter. The glucose concentrations for all volunteers ranged from 3.8 to 12.4 mM ($\sim$68–223 mg/dL). Raman spectra over the spectral range 1545–355 cm^{-1} were selected for data analysis. An average of 27 (461/17) spectra were obtained for each individual. Each spectrum was obtained with excitation power $\sim$300 mW in a 1 mm spot diameter on the forearm, and integration time equivalent to 3 minutes.

Spectra from each volunteer were analyzed using PLS with leave-one-out cross validation, with 8 factors retained for development of the regression vector. PLS with leave-one-out cross validation was first performed on each individual, a mean absolute error (MAE) of 7.8% (RMSECV $\sim$ 0.7 mM) and an R^2 of 0.83 were obtained. When the data from 9 volunteers were combined, the MAE was 12.8% with $R^2 \sim 0.7$, while combining all 17 volunteers gave MAE $\sim$ 16.9% (RMSECV $\sim$ 1.5 mM). In individual calibrations, the adequacy of the ratio of the number of PLS factors and that of the samples was a concern. However, in the combined data set, the grouping schemes involving 9 (244 spectra) and 17 (461 spectra) volunteers utilized 17 and

21 factors, respectively, which is acceptable [16]. Another encouraging piece of evidence was that multiple glucose spectral features were identified in the regression vectors, indicating that the calibration was at least partially based on glucose.

Since this was a feasibility study, the protocol did not include measurements on the volunteers over a number of days, and thus independent data was not obtained. Further, oral glucose tolerance test protocols are susceptible to correlation with the fluorescence background decay, which may enhance the apparent prediction results. Therefore, more studies, preferably involving glucose clamping performed on different days, are required.

12.8 Toward Prospective Application

The results from the *in vitro* and *in vivo* studies reviewed above are encouraging. They demonstrate the feasibility of building glucose-specific *in vivo* multivariate calibration models based on Raman spectroscopy. To bring this technique to fruition, prospective application of a calibration algorithm on independent data with clinically acceptable prediction results needs to be demonstrated. This requires advances in extracting glucose information without spurious correlations, and correcting for variations in subject tissue morphology and color. We have developed two new tools to address these issues. The first is a novel multivariate calibration technique with higher analyte specificity than present techniques, and is more robust against interferent co-variations and chance correlations. This technique, constrained regularization, is described in subsection 12.8.1. In addition, we have developed a new correction method to compensate for turbidity-induced sampling volume variations across sites and individuals. This method, intrinsic Raman spectroscopy, is discussed in subsection 12.8.2. Other considerations for successful *in vivo* studies, such as reference concentration accuracy and optimal collection site determination, will be discussed in the context of future directions.

12.8.1 Analyte-specific information extraction using hybrid calibration methods

Multivariate calibration methods are in general not analyte-specific. Calibration models are built based on correlations in the data, which may be due to the analyte or to systematic or spurious effects. One way to effectively boost the model specificity is by incorporating additional analyte-specific information, such as its pure spectrum. Hybrid methods merge additional spectral information with calibration data in an implicit calibration scheme. In the following, we present two of these methods developed in our laboratory.

12.8.2 Hybrid linear analysis (HLA)

Hybrid linear analysis was developed by Berger *et al*. [24]. In this method, analyte spectral contributions are first removed from the sample spectra by subtracting the pure spectrum according to reference concentration measurements. The resulting spectra are then analyzed by principal component analysis (PCA), and the significant principal components extracted. These principal components are then used as basis spectra to orthogonalize the pure analyte spectrum. This orthogonalization results in a ***b*** vector that is essentially the portion of the pure analyte spectrum that is orthogonal to all interferent spectra, akin to the net analyte signal.

HLA was implemented experimentally *in vitro* with a 3-analyte model composed of glucose, creatinine, and lactate. Significant improvement over PLS was obtained owing to incorporating the pure glucose spectrum in the model development. However, because HLA relies on subtracting the analyte spectrum from the calibration data, it is very sensitive to the accuracy of the spectral shape and intensity. For turbid samples with multiple analytes in which absorption and scattering can alter the analyte spectral features, we find that the performance of HLA is impaired. Motivated by advancing transcutaneous measurement of blood analytes *in vivo*, constrained regularization was developed as a more robust method against inaccuracies in the pure analyte spectra.

12.8.3 Constrained regularization (CR)

To understand constrained regularization, multivariate calibration can be viewed as an inverse problem. Given the inverse mixture model for a single analyte:

$$\mathbf{c} = \mathbf{S}^T\mathbf{b} \tag{12.4}$$

where capital boldface type denotes a matrix. The goal is to invert Eq. (12.4) and obtain a solution for **b**. Factor-based methods such as principal component regression (PCR) and PLS summarize the calibration data, $[\mathbf{S},\mathbf{c}]$, using a few principal components or loading vectors. In contrast, CR seeks a balance between model approximation error and noise propagation error by minimizing the cost function, Φ [49]:

$$\Phi(\Lambda,\mathbf{b}_0) = \left\|\mathbf{S}^T\mathbf{b}-\mathbf{c}\right\|^2 + \Lambda\left\|\mathbf{b}-\mathbf{b}_0\right\|^2 \tag{12.5}$$

with $\|\mathbf{e}\|$ the Euclidean norm (*i.e.*, magnitude) of **e**, and $\mathbf{b}_0$ a spectral constraint that introduces prior information about **b**. The first term on the right-hand side of Eq. (12.5) is the model approximation error, and the second term is the norm of the difference between the solution and the constraint, which controls the smoothness of the solution and its deviation from the constraint. If $\mathbf{b}_0$ is zero, the solution is the common regularized solution. For $\Lambda = 0$ the least squares solution is then obtained. In the other limit, in which Λ goes to infinity, the solution is simply $\mathbf{b} = \mathbf{b}_0$.

A reasonable choice for $\mathbf{b}_0$ is the spectrum of the analyte of interest, because that is the solution for **b** in the absence of noise and interferents. Another choice is the net analyte signal [28] calculated using all of the known pure analyte spectra. Such flexibility in the selection of $\mathbf{b}_0$ is possible because of the manner in which the constraint is incorporated into the calibration algorithm. For CR, the spectral constraint is included in a nonlinear fashion through minimizing Φ, and is thus termed a ”soft” constraint. On the other hand, there is little flexibility for methods such as HLA, in which the spectral constraint is algebraically subtracted from each sample spectrum before performing PCA. We term this type of constraint a ”hard” constraint.

Numerical simulations and *in vitro* experiments indicate that CR gives lower RMSEP than methods without prior information, such as PLS, and that it is less affected by analyte co-variations. We further demonstrated that CR is more robust than HLA when there are inaccuracies in the applied constraint, as often occurs in complex or turbid samples such as biological tissue [27]. The simulation and experimental results for RMSEP are summarized in Table 12.1, where all values were normalized to the PLS values. Glucose and creatinine were used as the analytes of interest with urea the major interferent. India ink and intralipid were used to provide turbidity with absorption and scattering levels similar to those of biological tissue.

TABLE 12.1: Comparison of RMSEP values for PLS, HLA, and CR. Glucose (G) and creatinine (C) are the analytes of interest. All RMSEP values are normalized to those of PLS. Note that all values of turbidity span the physiologically relevant range

	PLS	HLA		CR	
	G & C	G	C	G	C
Simulation	1	0.64	0.53	0.8	0.59
Clear samples	1	0.73	0.95	0.73	0.89
Turbid samples	1	1.15	1.12	0.8	0.88

An important lesson learned from this study is that there is a trade-off between maximizing prior information utilization, and robustness concerning the accuracy of this information. Multivariate calibration methods range from explicit methods with maximum use of prior information (*e.g.*, OLS, least robust when the calibration model is inaccurate), hybrid methods with an inflexible constraint (*e.g.*, HLA), hybrid methods with a flexible constraint (*e.g.*, CR), and implicit methods with no prior information (*e.g.*, PLS, most robust, but prone to be misled by spurious correlations). We believe that CR achieves the optimal balance between these ideals in practical situations.

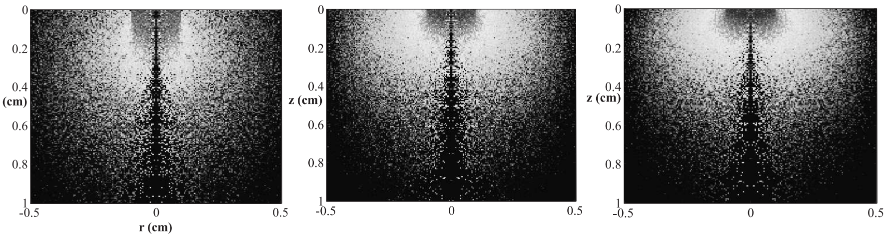

FIGURE 12.8: Turbidity-induced sampling volume variations simulated by the Monte Carlo method. Steady-state fluence rate owing to excitation for three turbidity-induced sampling volumes: (left) large; (middle) medium; (right) small sampling volume.

12.8.4 Sampling volume correction using intrinsic Raman spectroscopy

Sample variability is a critical issue in prospective application of a calibration model. For optical technologies, variations in tissue optical properties, particularly absorption (μ_a) and scattering (μ_s) coefficients, can distort the measured spectra. Figure 12.8 shows the results from Monte Carlo simulations, demonstrating that the effective sampling volume strongly depends on the optical properties of the medium. Note that the three values of turbidity shown are within the physiologically relevant range. If these turbidity-induced sampling volume variations are not corrected for, large errors will be introduced into the subsequent multivariate analysis.

This section provides an overview of techniques to correct turbidity-induced spectral and intensity distortions in fluorescence and Raman spectroscopy. Photon migration theory is employed to model diffuse reflectance, fluorescence and Raman scattering arising from turbid biological samples. Monte Carlo simulations are employed as an effective and statistically accurate tool to numerically model light propagation in turbid media. Using the photon migration model and Monte Carlo simulations, preliminary results of the use of intrinsic Raman spectroscopy to correct the effects of turbidity are presented. Details of these results will be published separately [50].

12.8.5 Corrections based on photon migration theory

Radiative transfer theory [21] provides an intensity based picture of light propagation in turbid media. However, the analytical solution to this integro-differential equation can be found only for very special conditions and approximations. The most extensively studied approximation is diffusion theory, which is used to model photons that undergo multiple scattering events [21]. Another very useful approximation, developed in our laboratory, is photon migration theory [51, 52]. This method employs probabilistic concepts to describe the scattering of light, and sets up a framework for an analytical expression relating the measured fluorescence to the intrinsic fluorescence, defined as the fluorescence as measured from a optically-thin slice of tissue free of the effects of scattering and absorption. This expression has been employed to recover turbidity-free fluorescence spectra from various types of tissue.

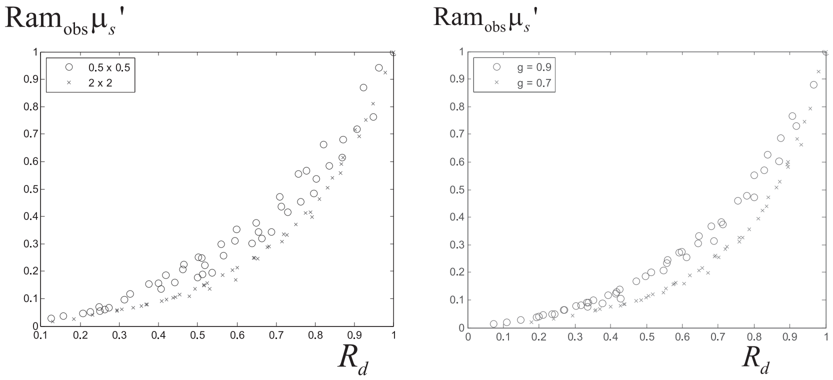

FIGURE 12.9: Ram$_{\text{obs}}\,\mu_s'$ versus R_d for samples of two different sizes (left) and for two different elastic scattering anisotropies (right). Note that all values of turbidity span the physiologically relevant range.

The correction facilitates interpretation of underlying fluorophores and consequently improves the accuracy of disease diagnosis [33, 53].

The same general principle should hold true for Raman spectroscopy, as well. Unlike fluorescence spectroscopy, spectral lineshape distortions caused by prominent absorbers is less of an issue in the NIR wavelength range. However, for quantitative analysis, the turbidity-induced sampling volume variations become significant. Monte Carlo simulations and experimental results show that the intrinsic Raman signal for arbitrary samples and collection geometries can be described by:

$$\text{Ram}_{\text{int}} = \mu_s' \frac{\text{Ram}_{\text{obs}}}{f(R_d)} \tag{12.6}$$

with $\mu_s' = \mu_s(1-g)$ and g the tissue anisotropy, and $f(R_d)$ the calibration factor with R_d the diffuse reflectance measured at the Raman wavelength. Fit parameters for $f(R_d)$ in Eq. (12.6) can be experimentally calibrated, and employed to obtain the intrinsic Raman signal in prospective spectra [54].

12.8.6 Intrinsic Raman spectroscopy (IRS)

Figure 12.9 (left) plots the product Ram$_{\text{obs}}\,\mu_s'$ versus R_d for two sample sizes, using the results from Monte Carlo simulations. A fit to this curve approximates the product of the intrinsic Raman signal and $f(R_d)$, and can be used to correct for sampling volume variations. We find that the curvature of $f(R_d)$ depends on the size of the sample. We also find that the curvature depends on g, as shown in Fig. 12.9 (right). Note that the higher spread for the samples of smaller size and larger anisotropy results from less returned photons, and thus a lower SNR.

IRS can be implemented either with numerical simulations or tissue phantoms. In either case, the implementation is done in two steps: calibration and application. In the calibration step, a set of samples with a wide range of turbidity values (within the physiological range) is studied, with a Raman scatterer of the same concentration inserted in each sample as a probe. Fig. 12.9 (left) can then be generated with fit parameters to obtain the functional form of $f(R_d)$. This function can then be used in the application step, in which the concentrations of the analytes are varied, to extract the intrinsic Raman signal from the measured Raman signal.

Note that to apply IRS one needs to know μ_s' for the samples. Extraction of optical properties has been studied by many researchers [55–58]. The majority of methods are based on diffusion theory or its variants. Our laboratory extracts optical properties from biological tissue routinely in other wavelength ranges, and a similar method can be employed for this purpose [56].

Intrinsic Raman spectroscopy is a novel technique. We look forward to incorporating it in our future *in vivo* studies. In addition, since CR and IRS both address the issues associated with non-analyte-specific correlations, yet from entirely different aspects, they can be applied in tandem. It will be interesting to study the resulting synergistic effects.

12.8.7 Other considerations and future directions

CR and IRS are two exciting topics for continued study. As mentioned earlier, there are several other areas also worthy of study. These are briefly discussed below.

12.8.7.1 Data quality

Reference concentrations greatly affect the performance of the calibration algorithm. In spectroscopic techniques such as Raman spectroscopy, a large portion of the collected glucose signal likely originates from the glucose molecules in the interstitial fluid (ISF). In addition, it is well known that the ISF glucose concentration in humans lags the plasma glucose concentration by 5 to 30 minutes [59]. As a result, use of plasma glucose as the reference concentration may introduce errors. Methods for extracting interstitial fluid for glucose reference measurements are being developed and should be incorporated in future studies [60, 61].

As discussed in subsection 12.4.2, the background and issues associated with it impose a limit on the SNR, and can influence multivariate calibration. Although the present background subtraction techniques address this issue, there is room for improvement. Thus, methods to reduce the background signal at its origin should be explored. One approach to separate fluorescence from Raman scattering is to study the spectrum at two closely-spaced excitation wavelengths and take the difference between them. It will be interesting to see if this technique can provide insight into the origin of the background signal and its variation.

12.8.7.2 Tissue morphology and skin heterogeneity

Sampling depth and sample positioning are critical for optimal collection of glucose-specific Raman scattered light. These may play a role in calibration transfer. In experiments, the effective sampling depth can be estimated from study of extracted optical properties, and therefore the correct distance between the sample and the collection optics can be determined for each measurement site. A study of morphological and layered structures at the sampling site with confocal microscopy and a properly designed human-instrument interface could shed light on the sampling depth and positioning. By knowing the exact sampling volume and its coverage of various skin morphological structures, it may be possible to estimate how much of the glucose-containing region (dermis in the two-layer model) is sampled.

12.9 Conclusion

Quantitative biological Raman spectroscopy is a powerful technique for non-invasive tissue analysis and analyte concentration measurements. From its early development with *in vitro* studies, *in vivo* studies have been realized with the aid of more advanced instrumentation and calibration algorithms. The *in vivo* studies performed to date have demonstrated the feasibility of obtaining glucose-specific multivariate calibration models. The next step in advancing the technology is to conduct prospective studies. Some of the issues to be addressed are enhancing analyte specificity, correcting for diversity across individuals, and issues relating to improved reference concentrations, and study of the role of tissue morphology and skin heterogeneity.

This chapter has reviewed recent developments in the first two categories by introducing constrained regularization and intrinsic Raman spectroscopy. With the aid of additional analyte spectral information, CR effectively improves analyte specificity. IRS, on the other hand, greatly reduces turbidity-induced sampling volume variations, one of the most challenging factors in multivariate calibration. These techniques will play a crucial role in prospective studies involving multiple sites/subjects/-days. We are currently planning a multi-subject and multiple-day *in vivo* study, first on dogs and then on human subjects. We believe that these new developments will enable us to demonstrate that quantitative Raman spectroscopy can accurately measure blood analyte concentrations prospectively.

Acknowledgements

Our own work was performed at the MIT Laser Biomedical Research Center and supported by NIH NCRR, Grant No. P41-RR0594, and a grant from Bayer Heath Care, LLC.

References

[1] R.L. McCreery, *Raman Spectroscopy for Chemical Analysis*, John Wiley & Sons, New York, 2000.

[2] C.V. Raman and K.S. Krishnan, “A new type of secondary radiation,” *Nature*, vol. 121, 1928, pp. 501–502.

[3] T. Hirschfeld and B. Chase, “Ft-Raman spectroscopy - development and justification,” *Appl. Spectr.*, vol. 40, 1986, pp.133–137.

[4] J.J. Baraga, M. S. Feld, and R. P. Rava, “Rapid near-infrared Raman spectroscopy of human tissue with a spectrograph and CCD detector,” *Appl. Spectr.*, vol. 46, 1992, pp. 187–190.

[5] H. Martens and T. Naes, *Multivariate Calibration*, John Wiley & Sons, New York, 1989.

[6] A.S. Haka, K.E. Shafer-Peltier, M. Fitzmaurice, J. Crowe, R.R. Dasari, and M. S. Feld, “Diagnosing breast cancer by using Raman spectroscopy,” *Proc. Nat. Acad. Sci. USA*, vol. 102, 2005, pp. 12371–12376.

[7] S.W.E. van de Poll, T. C. B. Schut, A. van den Laarse, and G.J. Puppels, “*In situ* investigation of the chemical composition of ceroid in human atherosclerosis by Raman spectroscopy,” *J. Raman Spectr.*, vol. 33, 2002, pp. 544–551.

[8] H.P. Buschman, J.T. Motz, G. Deinum, T.J. Romer, 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,” *Cardiovasc. Pathol.*, vol. 10, 2001, pp. 59–68.

[9] P.J. Caspers, G.W. Lucassen, E.A. Carter, H.A. Bruining, and G.J. Puppels, “*In vivo* confocal Raman microspectroscopy of the skin: Noninvasive determination of molecular concentration profiles,” *J. Invest. Dermatol.*, vol. 116, 2001, pp. 434–442.

[10] K.P. Peterson, P.E. Knudson, J. Chaiken, C.M. Peterson, W.F. Finney, X. Yang, and R.S. Weinstock, “Clinical use of noninvasive tissue-modulated Raman spectroscopic measurement of blood glucose in whole blood,” *Clin. Chem.*, vol.47, 2001, pp. A164–A165.

[11] J.N.Y. Qu, B.C. Wilson, and D. Suria “Concentration measurements of multiple analytes in human sera by near-infrared laser Raman spectroscopy,” *Appl. Opt.*, vol. 38, 1999, pp. 5491–5498.

[12] D. Rohleder, W. Kiefer, and W. Petrich, “Quantitative analysis of serum and serum ultrafiltrate by means of Raman spectroscopy,” *Analyst*, vol. 129, 2004, pp. 906–911.

[13] A.J. Berger, T.W. Koo, I. Itzkan, G. Horowitz, and M.S. Feld, “Multicomponent blood analysis by near-infrared Raman spectroscopy,” *Appl. Opt.*, 38, 1999, pp. 2916–2926.

[14] 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,” *Opt. Lett.*, vol. 27, 2002, pp. 2004–2006.

[15] J.L. Lambert, J.M. Morookian, S.J. Sirk, and M.S. Borchert, “Measurement of aqueous glucose in a model anterior chamber using Raman spectroscopy,” *J. Raman Spectr.*, vol. 33, 2002, pp. 524–529.

[16] A.M.K. Enejder, T.G. Scecina, J. Oh, M. Hunter, W.-C. Shih, S. Sasic, G.L. Horowitz, and M. S. Feld, “Raman spectroscopy for noninvasive glucose measurements,” *J.Biomed. Opt.*, vol. 10, 2005, 031114.

[17] J.N. Roe and B.R. Smoller “Bloodless glucose measurements,” *Critical Reviews in Therapeutic Drug Carrier Systems* 1998, 15, 199–241.

[18] A.C. Guyton and J.E. Hall, *Human Physiology and Mechanisms of disease*, 6th ed., Saunders, Philadelphia, 1997.

[19] P. Geladi and B.R. Kowalski, “Partial Least-Squares Regression - a Tutorial,” *Analyt. Chim. Acta*, vol. 185, 1986, pp. 1–17.

[20] B.R. Kowalski and A. Lorber, “Recent Advances in Multivariate Calibration,” *Abstracts of Papers of the American Chemical Society*, vol. 196, 1988, p.100–Anyl.

[21] A. Ishimaru, *Wave Propagation and Scattering in Random Media*, Academic Press, New York, 1978.

[22] L.H. Wang, S.L. Jacques, and L.Q. Zheng, “Mcml - Monte-Carlo modeling of light transport in multilayered tissues,” *Comp. Meth. Progr. Biomed.*, vol. 47, 1995, pp.131–146.

[23] M.A. Arnold, J.J. Burmeister, and G. W. Small, “Phantom glucose calibration models from simulated noninvasive human near-infrared spectra,” *Anal. Chem.*, vol. 70, 1998, pp. 1773–1781.

[24] A.J. Berger, T.W. Koo, I. Itzkan, and M.S. Feld, “An enhanced algorithm for linear multivariate calibration,” *Anal. Chem.*, vol. 70, 1998, pp. 623–627.

[25] D.M. Haaland and D.K. Melgaard, “New classical least-squares/partial least-squares hybrid algorithm for spectral analyses,” *Appl. Spectr.*, vol. 55, 2001, pp. 1–8.

[26] P.D. Wentzell, D.T. Andrews, and B.R. Kowalski, “Maximum likelihood multivariate calibration,” *Anal. Chem.*, vol. 69, 1997, pp. 2299–2311.

[27] W.-C. Shih, K.L. Bechtel, and M.S. Feld, “Constrained regularization: Hybrid method for multivariate calibration,” *Anal. Chem.*, vol. 79, 2007, pp. 234–239.

[28] A. Lorber, “Error propagation and figures of merit for quantification by solving matrix equations,” *Anal. Chem.*, vol. 58, 1986, pp. 1167–1172.

[29] A. Lorber, K. Faber, and B.R. Kowalski, “Net analyte signal calculation in multivariate calibration,” *Anal. Chem.*, vol. 69, 1997, pp.1620–1626.

[30] O.R. Scepanovic, K.L. Bechtel, A.S. Haka, W.-C. Shih, T.-W. Koo, A.J. Berger, and M.S. Feld, “Determination of uncertainty in parameters extracted from single spectroscopic measurements,” *J.Biomed. Opt.*, vol. 12, 2007, 064012.

[31] R. Gillies, G. Zonios, R.R. Anderson, and N. Kollias, “Fluorescence excitation spectroscopy provides information about human skin *in vivo*,” *J. Invest. Dermatol.*, vol. 115, 2000, pp. 704–707.

[32] E.L. Hull, M.N. Ediger, A.H.T. Unione, E.K. Deemer, M.L. Stroman, and J.W. Baynes, “Noninvasive, optical detection of diabetes: model studies with porcine skin,” *Opt. Exp.*, vol. 12, 2004, pp. 4496–4510.

[33] M.G. Muller, I. Georgakoudi, Q.G. Zhang, J. Wu, and M.S. Feld, “Intrinsic fluorescence spectroscopy in turbid media: disentangling effects of scattering and absorption,” *Appl. Opt.*, vol. 40, 2001, pp. 4633–4646.

[34] H. Zeng, C. MacAulay, D.I. McLean, B. Palcic, and H. Lui, “The dynamics of laser–induced changes in human skin autofluorescence – Experimental measurements and theoretical modeling,” *Photochem. Photobiol.*, vol. 68, 1998, pp. 227–236.

[35] A.J.L. Jongen and H.J.C.M. Sterenborg, “Mathematical description of photobleaching *in vivo* describing the influence of tissue optics on measured fluorescence signals,” *Phys. Med. Biol.*, vol. 42, 1997, pp. 1701–1716.

[36] I.B. Gornushkin, P.E. Eagan, A.B. Novikov, B.W. Smith, and J.D. Winefordner, “Automatic correction of continuum background in laser-induced breakdown and Raman spectrometry,” *Appl. Spectr.*, vol. 57, 2003,pp. 197–207.

[37] C.A. Lieber and A. Mahadevan-Jansen, “Automated method for subtraction of fluorescence from biological Raman spectra,” *Appl. Spectr.*, vol. 57, 2003, pp. 1363–1367.

[38] T. J. Vickers, R. E. Wambles, and C. K. Mann, “Curve fitting and linearity: Data processing in Raman spectroscopy,” *Appl. Spectr.* 2001, 55, 389–393.

[39] P.J. Caspers, G.W. Lucassen, and G.J. Puppels, “Combined *in vivo* confocal Raman spectroscopy and confocal microscopy of human skin,” *Biophys. J.*, vol. 85, 2003, pp. 572–580.

[40] T.-W. Koo, *Measurement of Blood Analytes in Turbid Biological Tissue Using Near-Infrared Raman Spectroscopy*, Massachusetts Institute of Technology, Cambridge, 2001.

[41] 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,” *Appl. Opt.*, vol. 43, 2004, pp. 542–554.

[42] Z.W. Huang, H.S. Zeng, I. Hamzavi, D.I. McLean, and H. Lui, “Rapid near-infrared Raman spectroscopy system for real-time *in vivo* skin measurements,” *Opt. Lett.*, vol. 26, 2001, pp. 1782–1784.

[43] J. Zhao, “Image curvature correction and cosmic removal for high-throughput dispersive Raman spectroscopy,” *Appl. Spectr.*, vol. 57, 2003, pp. 1368–1375.

[44] I. Pelletier, C. Pellerin, D.B. Chase, and J.F. Rabolt, “New developments in planar array infrared spectroscopy,” *Appl. Spectr.* vol. 59, 2005, pp. 156–163.

[45] C.H. Spiegelman, M.J. McShane, M.J. Goetz, M. Motamedi, Q.L. Yue, and G. L. Cotè, “Theoretical justification of wavelength selection in PLS calibration development of a new algorithm,” *Anal. Chem.*, vol. 70, 1998, pp. 35–44.

[46] A.J. Berger, Y. Wang, D.M. Sammeth, I. Itzkan, K. Kneipp, and M.S. Feld, “Aqueous Dissolved-Gas Measurements Using near-Infrared Raman-Spectroscopy,” *Appl. Spectr.*, vol. 49, 1995, pp. 1164–1169.

[47] A.J. Berger, Y. Wang, and M.S. Feld, “Rapid, noninvasive concentration measurements of aqueous biological analytes by near-infrared Raman spectroscopy,” *Appl. Opt.*, vol.35, 1996, pp. 209–212.

[48] A.J. Berger, I. Itzkan, and M.S. Feld, “Feasibility of measuring blood glucose concentration by near-infrared Raman spectroscopy,” *Spectrochim. Acta Part a-Molec. Biomolec. Spectr.*, vol.53, 1997, pp. 287-292.

[49] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Institute of Physics Pub., Bristol, UK , Philadelphia, PA, 1998.

[50] W.-C. Shih, K.L. Bechtel, and M.S. Feld, “Intrinsic Raman spectroscopy for quantitative biological spectroscopy, Part I: Theory and simulations,” in preparation.

[51] J. Wu, F. Partovi, M.S. Field, and R.P. Rava, “Diffuse Reflectance from Turbid Media - an Analytical Model of Photon Migration,” *Appl. Opt.*, 32, 1993, pp. 1115–1121.

[52] J. Wu, M.S. Feld, and R.P. Rava, “Analytical model for extracting Intrinsic fluorescence in turbid media,” *Appl. Opt.*, vol. 32, 1993, pp. 3585–3595.

[53] Q.G. Zhang, M.G. Muller, J. Wu, and M.S. Feld, “Turbidity-free fluorescence spectroscopy of biological tissue,” *Opt. Lett.*, vol. 25, 2000, pp. 1451–1453.

[54] K.L. Bechtel, W.-C. Shih, and M.S. Feld, “Intrinsic Raman spectroscopy for quantitative biological spectroscopy, Part II: Experiments”, in preparation.

[55] T.J. Farrell, M.S. Patterson, and B. Wilson, “A diffusion-theory model of spatially resolved, steady-state diffuse reflectance for the noninvasive determina-

tion of tissue optical-properties *in vivo*," *Med. Phys.*, vol. 19, 1992, pp. 879–888.

[56] G. Zonios, L.T. Perelman, V.M. Backman, R. Manoharan, M. Fitzmaurice, J. Van Dam, and M.S. Feld, "Diffuse reflectance spectroscopy of human adenomatous colon polyps *in vivo*," *Appl. Opt.*, vol.38, 1999, pp. 6628–6637.

[57] R.M.P. Doornbos, R. Lang, M.C. Aalders, F.W. Cross, and H.J.C.M. Sterenborg, "The determination of *in vivo* human tissue optical properties and absolute chromophore concentrations using spatially resolved steady-state diffuse reflectance spectroscopy," *Phys. Med. Biol.*, vol. 44, 1999, pp. 967–981.

[58] M.G. Nichols, E.L. Hull, and T.H. Foster, "Design and testing of a white-light, steady-state diffuse reflectance spectrometer for determination of optical properties of highly scattering systems," *Appl. Opt.*, vol. 36, 1997, pp. 93–104.

[59] M.S. Boyne, D.M. Silver, J. Kaplan, and C.D. Saudek, "Timing of changes in interstitial and venous blood glucose measured with a continuous subcutaneous glucose sensor," *Diabetes*, vol. 52, 2003, pp. 2790–2794.

[60] D.G. Maggs, R. Jacob, F. Rife, R. Lange, P. Leone, M.J. During, W.V. Tamborlane, and R.S. Sherwin, "Interstitial Fluid Concentrations of Glycerol, Glucose, and Amino-Acids in Human Quadricep Muscle and Adipose-Tissue - Evidence for Significant Lipolysis in Skeletal-Muscle," *J. Clin. Invest.*, vol. 96, 1995, pp. 370–377.

[61] J.P. Bantle and W. Thomas, "Glucose measurement in patients with diabetes mellitus with dermal interstitial fluid," *J. Lab. Clin. Med.*, vol. 130, 1997, pp. 436–441.