

Non-invasive Glucose Sensing with Raman Spectroscopy

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

1.	Introduction to Raman spectroscopy	3
2.	Biological considerations for Raman spectroscopy	4
2.1	Using near infrared radiation	4
2.2	Background signal in biological Raman spectra.....	6
2.3	Heterogeneities in human skin.....	7
3.	Quantitative considerations for Raman spectroscopy.....	7
3.1	Minimum detection error analysis	7
3.2	Multivariate calibration.....	9
4.	Instrumentation	10
4.1	Excitation light source	10
4.2	Light delivery.....	11
4.3	Light collection	12
4.4	Light transport.....	12
4.5	Spectrograph and detector.....	13
5.	Data Pre-Processing	14
5.1	Image curvature correction	14
5.2	Spectral range selection	17
5.3	Cosmic ray removal	17
5.4	Background subtraction	18
5.5	Random noise rejection and suppression.....	18
5.6	White light correction and wavelength calibration.....	19
6.	<i>In vitro</i> studies.....	19
6.1	Aqueous humor	20
6.2	Blood serum	21
6.3	Whole blood.....	23
7.	<i>In vivo</i> studies	23
7.1	Tissue modulation approach	23
7.2	Direct approach.....	25
8.	Toward prospective application.....	26
8.1	Analyte-specific information extraction using hybrid calibration methods	27
8.1.1	Hybrid linear analysis (HLA)	27
8.1.2	Constrained regularization (CR).....	28
8.2	Sampling volume correction using intrinsic Raman spectroscopy	30
8.2.1	Optical properties biological tissue.....	30
8.2.2	Corrections based on photon migration theory.....	32
8.2.3	Monte Carlo method	34
8.2.4	Intrinsic Raman spectroscopy (IRS)	35
8.3	Other considerations and future directions	36

9. Conclusion	38
---------------------	----

1. Introduction to Raman spectroscopy

Light that is scattered from a molecule is primarily elastically scattered, that is, the incident and the scattered photons have the same energy. A small probability exists, however, that a photon is scattered inelastically, resulting in either a net gain or loss of energy of the scattered photon. This inelastic scattering, discovered by Raman and Krishna,¹ allows fundamental molecular vibrational transitions to be measured at any excitation wavelength.

Raman scattering is a coherent one-step process in which one photon is exchanged for another through interaction with a molecule. Schematically, the Raman process is depicted as a molecule in an initial vibrational state proceeding to a higher or lower vibrational state through excitation to a “virtual state,” with simultaneous scattering of a new photon from this state. The difference in energy between the incident and scattered photon is equal to the energy difference between the initial and final vibrational states of the molecule. A loss in photon energy is termed Stokes-Raman scattering and a gain in photon energy is termed anti-Stokes-Raman scattering.² These processes are depicted in Figure 1.

Not all vibrational transitions can be accessed by Raman scattering. Raman-active transitions are those associated with a change in polarizability of the molecule. In classical terms, this can be viewed as a perturbation of the electron cloud of the molecule.

A Raman spectrum is a plot of scattered light intensity versus energy shift (also called Raman shift) reported in wavenumbers (cm^{-1}). An example spectrum of aqueous glucose is shown in Figure 2. To convert from a wavenumber shift to wavelength, the incident wavelength must be known. For example, a 600 cm^{-1} Raman shift occurs at 873.5 nm if

the excitation wavelength is 830 nm ($1 \times 10^{-7} / (1/830 \times 10^{-7} \text{ cm}^{-1} - 600 \text{ cm}^{-1}) = 873.5 \text{ nm}$) or at 823.8 nm if the excitation wavelength is 785 nm.

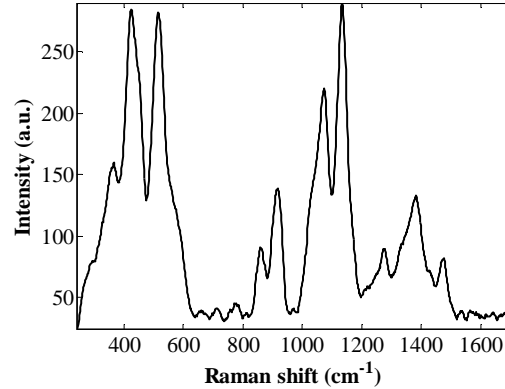
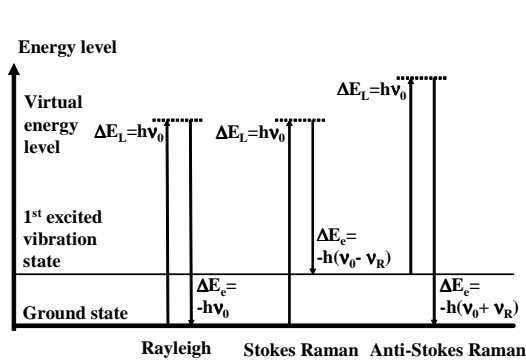


Figure 1. Energy diagram for Rayleigh, Stokes Raman, and anti-Stokes Raman scattering.

Figure 2. A Raman spectrum consists of scattered intensity plotted vs. energy. This figure uses glucose water solution measured in a quartz cuvette as an example.

In this chapter we focus on non-resonant spontaneous Raman scattering. A special case of Raman scattering, surface-enhanced Raman spectroscopy (SERS), is discussed in Chapter X.

2. Biological considerations for Raman spectroscopy

2.1 Using near infrared radiation

Raman shifts are independent of excitation wavelength and thus Raman spectroscopy offers the flexibility to select a suitable excitation wavelength for a specific application. The choice of NIR excitation for probing biological tissue is justified by three advantageous features: low-energy optical radiation, deep penetration, and reduced background fluorescence. Light in the NIR region is non-ionizing and therefore does not

pose the same exposure risk as X-ray radiation. Additionally, NIR light penetrates relatively deep into the tissue, on the order of mm – cm in some spectral windows. These depths are possible due to reduced elastic scattering, which decreases at longer wavelengths, and the lack of significant absorption bands in this region. Fluorescence is also much lower in the NIR region as compared to shorter wavelengths, thus allowing the less intense Raman bands to be resolved.

Figure 3 illustrates the absorption spectra of major endogenous tissue absorbers, namely, 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,” in which a group of minima exists, is outlined.

A final consideration for the selection of excitation wavelength in Raman spectroscopy is the efficiency of the silicon-cased charge coupled device (CCD) detector. Due to silicon absorption, CCD detectors are prohibitively inefficient above 1000 nm. As a result, 785 nm or, more recently, 830 nm are often chosen as the excitation wavelength to fully exploit the “diagnostic window” while retaining an acceptable quantum efficiency detector.

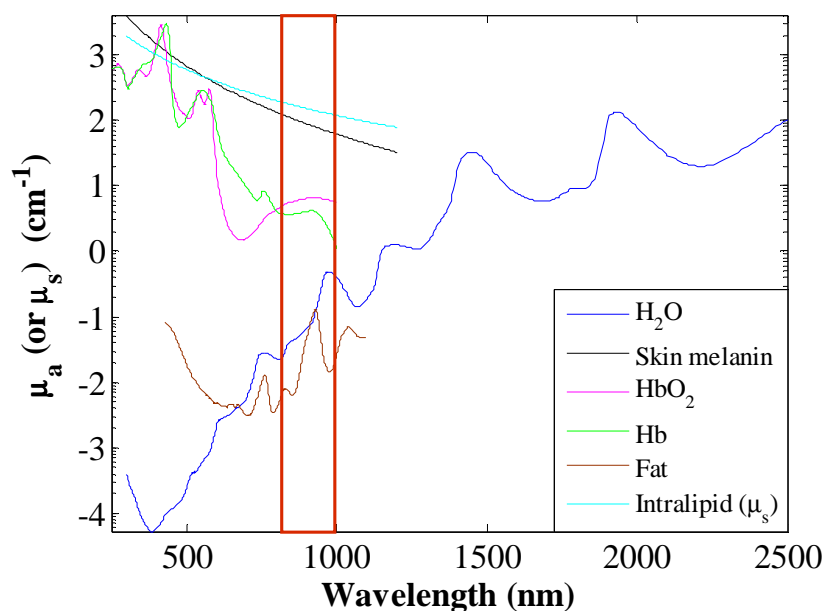


Figure 3. Absorption spectra of water, skin melanin, hemoglobin, and fat. Also shown is the scattering spectrum of 10% intralipid, a lipid emulsion often used to simulate tissue scattering. Data are obtained from <http://omlc.org.edu/spectra/index.html>.

2.2 Background signal in biological Raman spectra

Although greatly reduced in intensity as compared to UV-visible excitation, NIR Raman spectra of biological samples are often accompanied by a strong background, generally attributed to fluorescence. Macromolecules such as proteins and lipids are thought to contribute to the fluorescence background.³ Although Raman bands are clearly distinguished above the background, its presence results in higher shot noise and therefore decreases the signal-to-noise ratio. Further, the background decreases as a function of time with accompanying spectral variation. This variation interferes with the multivariate analysis. Thus, it is desirable to either reduce the background during data collection or remove it in pre-processing without introducing artifacts. Most background removal methods in the literature are based on low-order polynomial fitting and subsequent subtraction. Many researchers have found that a fifth-order polynomial best

approximates the shape of the background.⁴⁻⁷ However, because of the inevitable introduction of spectral artifacts, some researchers have found that removing the background does not improve calibration results obtained from multivariate analysis.⁸

2.3 Heterogeneities in human skin

Uniform analyte distribution is often a good assumption for liquid samples such as blood serum or even whole blood if stirring is continuous. For biological tissue, human skin in particular, heterogeneity is a major factor. Detailed morphological structures and molecular constituents of skin have been studied using confocal Raman spectroscopy.⁹

The skin is a layered system with two principle layers: epidermis and dermis. The epidermis is the outmost layer of skin and itself consists of multiple layers such as the stratum corneum, stratum lucidum, and stratum granulosum. The major constituent of human epidermis is keratin, comprising approximately 65% of the stratum corneum. The dermis is also a layered tissue composed of mainly collagen and elastin. Blood capillaries are present in the dermis and thus this region is targeted for optical analysis. However, it has been suggested that the majority of the glucose molecules sampled by a non-invasive optical technique are present in the interstitial fluid, which is usually found at the epidermis-dermis interface.¹⁰

3. Quantitative considerations for Raman spectroscopy

3.1 Minimum detection error analysis

If all component spectra in a mixture sample are known, the minimum detection error can be calculated via a simple formula derived by Koo *et al.*:^{10, 11}

$$\Delta c = \frac{\sigma}{s_k} \text{ of } f_k. \quad (1)$$

The first factor on the right hand side, σ , describes the noise in the measured spectrum and the second factor, s_k , describes the signal strength, calculated as the norm of the k^{th} model component. The last factor, olf_k , is termed the “overlap factor” and can take on values between 1 and ∞ .

The overlap factor indicates the amount of non-orthogonality (overlap) between the k^{th} model component and the other model components. Mathematically, the overlap factor is equal to the inverse of the correlation coefficient between the k^{th} component spectrum, \mathbf{s}_k , and the OLS regression vector, \mathbf{b}_{OLS} :

$$\text{olf}_k = \frac{1}{\text{corr}(\mathbf{b}_{\text{OLS}}, \mathbf{s}_k)}. \quad (2)$$

The OLS regression vector, also called the net analyte signal, is the part of the k^{th} component spectrum that is orthogonal to all interferences. It is equivalent to the k^{th} component spectrum when no interferences exist.

Correlation between two vectors is calculated by:

$$\text{corr}(u, v) = \frac{\sum_{i=1}^n (u_i - \bar{u})(v_i - \bar{v})}{\sqrt{\sum_{i=1}^n (u_i - \bar{u})^2 \sum_{i=1}^n (v_i - \bar{v})^2}}. \quad (3)$$

In the absence of interferences, the correlation coefficient is equal to 1 and therefore $\text{olf}_k = 1$. In this case, the minimum detection error is defined solely on the basis of signal-to-noise considerations. When interferences exist, the correlation coefficient is always smaller than 1 and therefore olf_k is always larger than 1. The minimum detection error approaches infinity when there is complete overlap and the analyte signal is indistinguishable.

To estimate the overlap factor for Raman measurements of glucose in skin, we have used a 10-component skin-mimicking model. Beginning with the spectrum of glucose, spectra of other constituents with strong Raman signals including collagen type I, keratin, triolein, actin, collagen type III, cholesterol, phosphatidylcholine, hemoglobin, and water were added one at a time to increase the model complexity. At each addition the correlation between \mathbf{b}_{OLS} and the glucose spectrum was calculated and ranges from 1 to 0.73 as shown in Figure 4. Therefore the overlap factor for glucose and skin is estimated at 1.4.

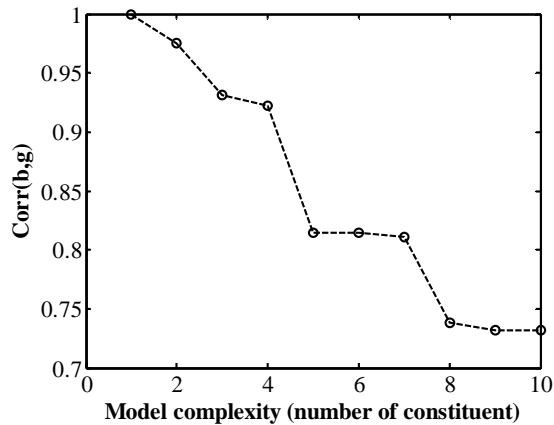


Figure 4 Correlation between the OLS regression vector and the glucose spectrum versus model complexity.

molecular specificity, spectral overlap is inevitable with the presence of multiple constituents. Further, the glucose Raman signal is only 0.3% of the total skin Raman signal.^{12, 13} Taken into consideration with the varying fluorescence background and random noise, it is not feasible to quantify the glucose signal by recording the skin Raman spectrum at only a few wavelengths. For quantitative analysis, multivariate techniques, which utilize the full-range spectra, are employed. In multivariate calibration, a set of calibration spectra and the associated glucose concentrations are used to calculate a regression vector. This regression vector, or \mathbf{b} vector, can be applied to a future

The high molecular specificity of Raman spectra results in less spectral overlap thus enabling the detection of low signal strength components such as glucose.

3.2 Multivariate calibration

As discussed previously, although Raman spectroscopy provides good

independent spectrum with unknown glucose content to extract the concentration.¹⁴⁻¹⁶ An introduction to multivariate techniques is in Chapter X, section Y.

4. Instrumentation

As discussed previously, background fluorescence impedes observation of Raman signal from biological tissue using UV-visible excitation wavelengths. To overcome this limitation, NIR excitation was employed with Fourier-transform spectrometers in the late 1980s.¹⁷ With the advent of high quantum efficiency CCD detectors and holographic diffractive optical elements, researchers have increasingly employed CCD-based dispersive spectrometers.^{3, 18-24} 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 better 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.

4.1 Excitation light source

Laser excitation at one of two wavelengths, 785 and 830 nm, is most common. The tradeoff lies in 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. In other embodiments, argon-ion laser pumped titanium-sapphire lasers are used extensively. The titanium-sapphire laser can provide higher power output

with broader wavelength tunability, but is bulkier and more expensive to maintain than diode lasers.

Because Raman scattering occurs at the same energy shift regardless of the excitation wavelength, narrowband 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 obscure low-wavenumber Raman bands. This problem is most severe in high power diode lasers and a holographic bandpass or interference laser line filter with attenuation greater than 6 optical density (OD) is usually required. Lastly, for quantitative measurements a photodiode is often needed to monitor the laser intensity to correct for fluctuations.

4.2 Light delivery

The filtered laser light can be delivered to the sample either through free-space or through an optical fiber. In the 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 is an important consideration 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 suppress the elastically scattered light entering the

collection fibers.²⁵ Commercial probes are now available and they offer ruggedness and easy access to samples with various special or geometrical constraints.

4.3 Light collection

As Raman scattering is a weak process, photons are precious and high collection efficiency is desired for a higher signal-to-noise ratio. Specialized optics such as Cassegrain microscope objectives and non-imaging paraboloidal mirrors have been employed to increase both the collection spot size and the effective numerical aperture of the optical system.¹⁰

The majority of photons that exit the air-sample interface are elastically scattered and remain at the original laser wavelength. This light must be properly attenuated or it will saturate the entire CCD detector. Holographic notch filters are extensively employed for this purpose and can attenuate the elastically scattered light to greater than 6OD, while passing the Raman photons with greater than 90% efficiency. However, notch filters are very sensitive to the incident angle of light and thus provides less attenuation to off-axis light. In some instances, the size 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,²⁶ 90 degree collection geometry,³ and a hole in the collection mirror have been realized to reduce its effect.²⁷

4.4 Light transport

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 rectangular shape of the entrance slit of a spectrograph, the originally round-shape of the collected light can be

relayed by an optical fiber bundle with the receiving end arranged into a round shape and the exiting end arranged linearly.²²

4.5 Spectrograph and detector

In dispersive spectrographs for Raman spectroscopy, transmission holographic gratings are often used for 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 $\sim 1\mu\text{m}$. 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 in pre-processing (described below).

As an example of these design considerations, Figure 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" holographic notch filter to reduce elastically scattered light. The Raman photons are 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 optic elements. The back-thinned, deep depletion, liquid nitrogen-cooled CCD is 1300x1340 pixels, height-matched to the fiber

bundle slit. This instrument was specifically designed for high sensitivity measurements in turbid media.

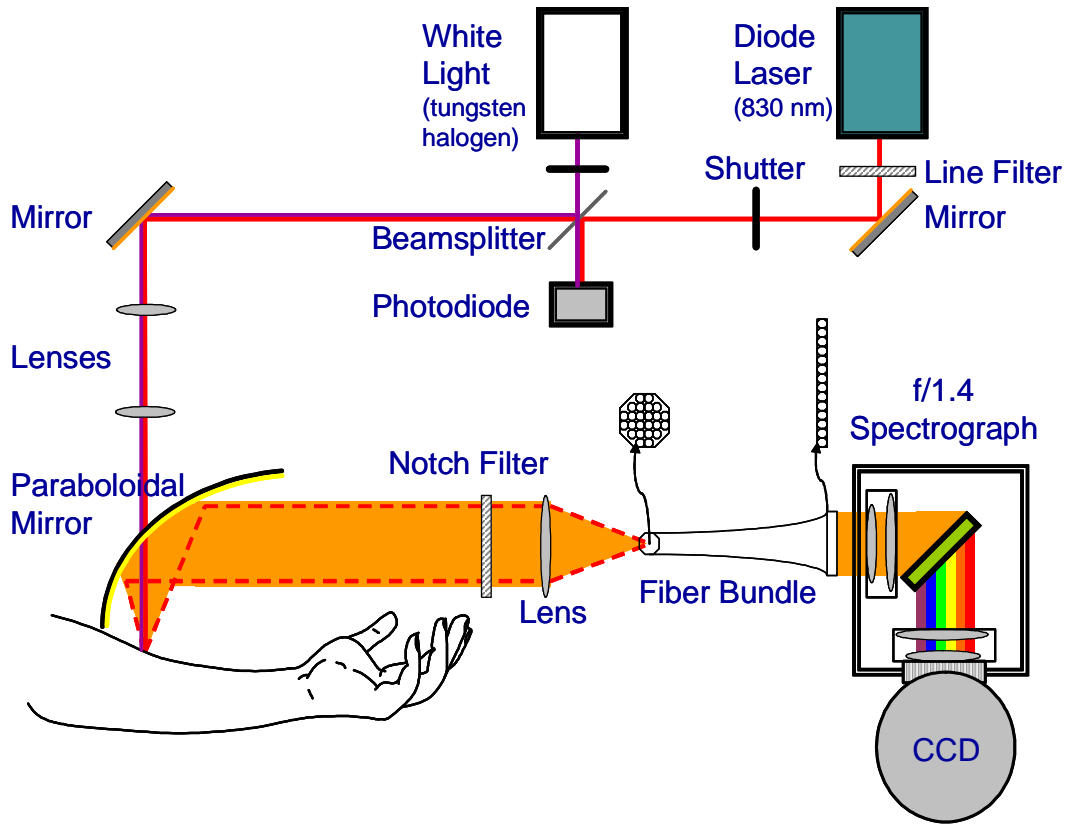


Figure 5 Schematic of a free space Raman instrument for noninvasive glucose measurements used at the MIT Spectroscopy Laboratory.

5. 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 for researchers to thoroughly document the exact steps taken.

Frequently employed procedures are described in the following.

5.1 Image curvature correction

Increase of usable detector area is an effective way to improve light throughput in Raman spectroscopy employing multi-channel dispersive spectrographs. However, owing to out-

of-plane diffraction the entrance slit image appears curved.²⁸ Direct vertical binning of detector pixels without correcting the curvature results in degraded spectral resolution.

Various hardware approaches, such as employing curved slits^{26, 28} or convex spherical gratings, have been implemented.²⁹ In the curved slit approaches, fiber bundles have been 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 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 (~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 our past research, we have developed a software method involving using a highly Raman active reference material to provide a sharp image on the CCD.⁸ 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.²⁸ 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.³⁰ In this section we present a similar method that was 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 calibration, 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, as depicted by the arrows in Figure 6.

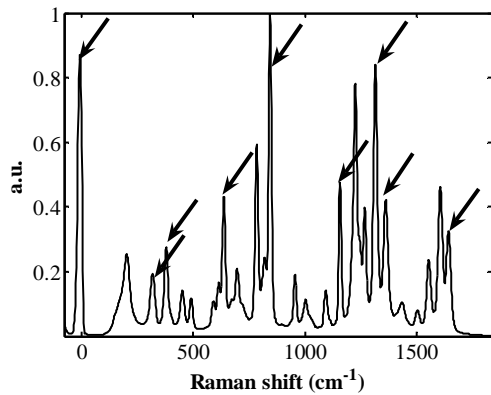


Figure 6 Raman spectrum of acetaminophen powder, used as the reference material in the calibration step. Nine prominent peaks used as separation boundaries are indicated by arrows.

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.

Significant improvement is observed from the pixel shift method to the curvature mapping method, especially toward either side of the CCD, as can be seen by comparing Figure 7(c) and Figure 7(e). The overall linewidth

reduction in 14 prominent peaks is 7% (FWHM). Such improvement is significant considering that the equivalent slit width is $\sim 360 \mu\text{m}$. If a narrower slit is employed for better spectral resolution, the overall linewidth reduction will be 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.

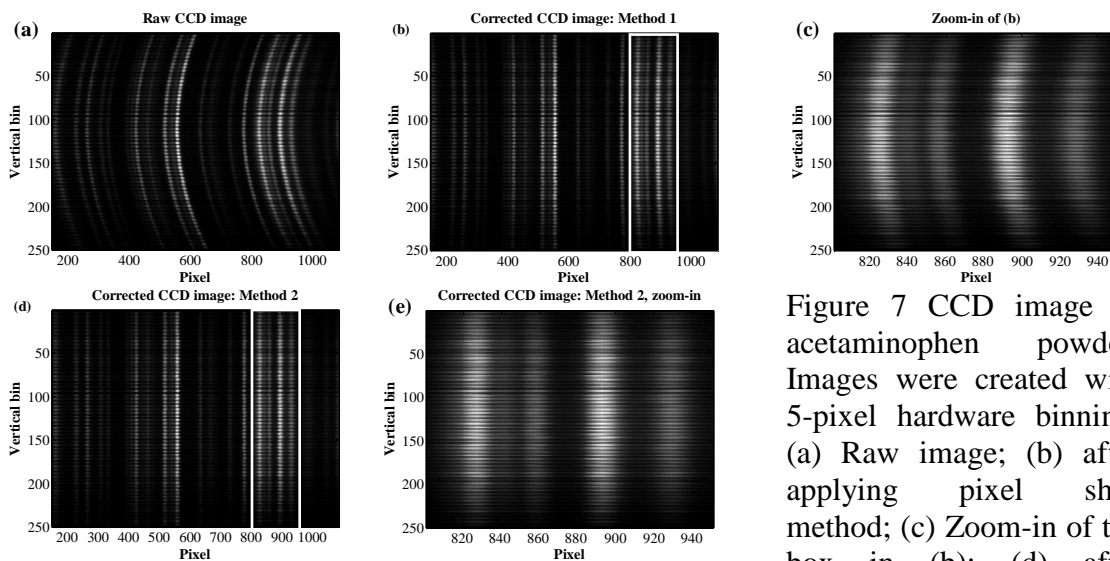


Figure 7 CCD image of acetaminophen powder. Images were created with 5-pixel hardware binning. (a) Raw image; (b) after applying pixel shift method; (c) Zoom-in of the box in (b); (d) after applying curvature mapping method; (e) Zoom-in of the box in (d).

5.2 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 $300\text{-}1700 \text{ cm}^{-1}$ is often chosen for analysis.

5.3 Cosmic ray removal

Cosmic rays hit the CCD array at random times with arbitrary intensity, 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.

5.4 Background subtraction

As mentioned in the biological considerations section, the background signal in Raman spectra is one of the limiting factors in determining the detection limit. Background removal techniques only approximate the shape of the background and therefore improvement in further quantitative analysis is often limited. However, for qualitative analysis, background-removed spectra provide better interpretation of the underlying constituents.

5.5 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 random noise can be estimated by the square root of the measured intensity. The most effective way to increase the signal-to-noise ratio (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 timescale offers no extra benefit as other errors begin to dominate performance.¹⁰ 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 an optimal number exists for binned pixels.¹⁰ However, the drawback to this method is degradation in spectral resolution. More commonly employed are Savitzky-Golay smoothing algorithms, which retains the data length.

5.6 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 to a spectrum from a calibrated light source, for example a calibrated tungsten-halogen lamp measured under identical conditions. Combinatorial spectral responses of the optical components, the diffraction grating, and the CCD camera can be effectively removed, and thus reveal 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, allowing for comparison of Raman features across instruments and time.

6. *In vitro* studies

In the following sections, we review the application of Raman spectroscopy to glucose sensing *in vitro*. *In vitro* studies have been performed using human aqueous humor, filtered and unfiltered human blood serum, and human whole blood, with promising results. Results in measurement accuracy are reported in root mean squared error values, with RMSECV for cross-validated, and RMSEP for predicted values. The reader is referred to the introductory chapter for discussion on these statistics.

6.1 Aqueous humor

Lambert *et al.* have explored the use of Raman spectroscopic measurements of glucose present in the aqueous humor of the eye.^{20, 21, 31} This is undoubtedly an excellent portal for optical measurements with potential advantages such as easy access and less complex fluid composition. In spite of these advantages, a spectroscopic measurement in the eye carries the risk of injury if the probing light is too intense. Therefore, dosimetry and a fool-proof light delivery method are important concerns for *in vivo* human studies.

Recently they demonstrated *in vitro* predictive capability of Raman spectroscopic measurements using a PLS calibration model derived from an artificial model.^{21, 31} Human aqueous humor (HAH) was used as the *in vitro* sample for prediction and artificial aqueous humor (AAH) was used to construct the calibration model. In the AAH model, five analytes including glucose at physiological concentrations were designed to vary independently with little correlation between any two analytes. The main advantage of using an AAH model is to break the glucose-lactate correlation present in HAH (correlation coefficient $r \sim 0.4$). The sample was placed in a contact lens for measurement by a Raman instrument using 785-nm excitation and a microscope objective with 180° geometry for Raman signal collection. Each spectrum was obtained with excitation power ~ 100 mW and integration time ~ 150 seconds.

They obtained an RMSEP of approximately 1-1.5 mM with $R^2 \sim 0.99$. Glucose spectral features were clearly observed in the second PLS factor, further supporting the calibration accuracy.³¹ They pointed out future directions such as focusing on demonstrating safety and efficacy in humans and determining the relationship between blood glucose and aqueous humor glucose.

6.2 Blood serum

Unprocessed samples

Our laboratory began investigating noninvasive blood analysis using Raman spectroscopy in the mid 1990's.³²⁻³⁴ The first biological sample study was conducted on serum and whole blood samples from 69 patients over a seven-week period.²³ 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 ~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 signal from the high turbidity. Glucose spectral features were identified in both the PLS weighting vector and the **b** vector, supporting that the calibration model was based on glucose.

With ultrafiltration

Qu *et al.*³ described the use of Raman spectroscopy for noninvasive glucose measurements in human serum samples after ultrafiltration, a process to remove macromolecules. Ultrafiltration can effectively eliminate Raman signals from large protein molecules that dominate unfiltered serum samples, thus significantly improving the signal-to-noise ratio and therefore the detection limit. Nevertheless, it is time consuming and requires extra sample preparation.

The experimental setup employed 785-nm excitation with a 90° collection geometry. Each spectrum was obtained with excitation power ~300 mW and integration time equivalent to 2.5 minutes. Because filtered serum is nearly transparent at 785 nm, excitation of Raman scattering is effectively along the entire laser path, creating a line-source in the cuvette. Thus the authors surmise that better collection efficiency could be obtained with optics designed specifically for this type of source, as opposed to the standard spherical lens they employed.

Regardless of potential collection efficiency improvements, they obtained an RMSEP of 0.38 mM. However, the PLS calibration model was obtained using 30 samples with 12 factors retained for the development of the regression vector. Without reported evidence of glucose spectral features it is difficult to determine whether the data was overfit. The model was applied to 24 samples that were not in the calibration set, giving some justification to the analysis.

Rohleder *et al.*³⁵ described measurement of glucose in both serum and filtered serum from 247 blood donors. A commercial spectrometer was used to acquire the spectra with 785-nm excitation and a double holographic grating covering a wavelength range of 785-1082 nm. Spectra were obtained with excitation power ~200 mW at the sample and integration time ~300 s. PLS calibration models were generated based on 148 samples and employed to predict the concentrations of the remaining 99 samples. They obtained an RMSEP of 0.95 ($R^2 \sim 0.97$) and 0.34 ($R^2 \sim 0.996$) mM in unfiltered and filtered serum samples, respectively.

6.3 Whole blood

The main difficulty for measurement in whole blood as compared to serum is the much higher absorption and scattering of whole blood attributed to the presence of hemoglobin and red blood cells, respectively. The combinatorial effect is a ~4X decrease in collected analyte Raman signal.

A subsequent whole blood study by Enejder *et al.*²² in our laboratory confirmed this hypothesis and they were able to demonstrate the feasibility of measuring multiple analytes in 31 whole blood samples with laser intensity and integration time similar to the previously mentioned serum study.²³ A 4X increase in Raman signal collection was achieved by employing a paraboloidal mirror and a shape-transforming fiber bundle for better collection efficiency, as depicted in Figure 5. PLS leave-one-out cross validation was performed and an RMSECV of 1.2 mM was obtained. The number of PLS factors compared to the number of samples raises the concern of overfitting. However, glucose spectral features were identified in the regression vector, providing more confidence that the model was based on glucose.

7. In vivo studies

7.1 Tissue modulation approach

Chaiken *et al.*^{24, 36, 37} proposed using Raman spectroscopy to measure glucose *in vivo* with a technique called “tissue modulation,” *i.e.*, continuously press/unpress the measurement site with a mechanical apparatus. The basic principle is that during the “press” phase, blood is expelled from the measuring site and thus the spectrum is considered as nearly devoid of blood. During the “unpress” phase, the spectrum is considered to be a sum of both blood and other tissue constituents.

In a recent report,²⁴ the difference spectrum between pressed and unpressed phases was interpreted as the whole blood spectrum. Each spectrum was obtained with excitation power ~ 31 mW and integration time ~ 100 seconds. Apparent glucose signal and blood volume factor were extracted by summing over $686\text{-}375\text{ cm}^{-1}$ and $1800\text{-}1000\text{ cm}^{-1}$ in the difference spectrum, respectively. Integrated normalized unit (INU) was then defined as the ratio of the apparent glucose signal to the blood volume factor. They claimed that $375\text{-}686\text{ cm}^{-1}$ contains the most glucose information and $1800\text{-}1000\text{ cm}^{-1}$ contains mostly fluorescence plus Raman signal from other tissue constituents, indicative of blood volume.

18 spectroscopic samples paired with fingerstick reference measurements were collected from an individual over two days with a time-randomized protocol. A calibration model was built by fitting a line through the plot of INU versus reference glucose concentration. The model was then applied to 31 samples collected over the next 14 weeks (7 additional samples from the same individual and 24 from different individuals). After rejecting 11 outliers, they obtained a correlation coefficient (r) of 0.8 and a standard deviation of ~ 1.2 mM.

Their work presents an interesting idea to isolate the glucose-containing blood spectrum by tissue modulation. The result suggests that the INU not only correlates to the reference glucose concentration, but also hemoglobin, which raises the concern of specificity. This technique can potentially be useful if several issues can be addressed. First, summation over $686\text{-}375\text{ cm}^{-1}$ obviously includes contributions from interferents and it is not clear how much error that introduces. In addition, hemoglobin is not the only substance that fluoresces and thus it is unclear why the calculated blood volume

factor could be representative of actual blood volume. Further, it is likely that most glucose Raman signal measured from skin originates from glucose molecules in the interstitial fluid. It is therefore unclear if the results were indeed based on blood glucose as claimed by the authors.

7.2 Direct approach

In our laboratory, Enejder *et al.* conducted a transcutaneous study on 17 non-diabetic volunteers using a version of the instrument depicted in Figure 5.⁸ Spectra were collected from the forearm of human volunteers in conjunction with an oral glucose tolerance test protocol, involving 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 fingerstick blood samples and subsequently analyzed by a Hemocue device. The glucose concentrations for all volunteers ranged from 3.8 to 12.4 mM (~68-223 mg/dL).

Raman spectra in the range 1545-355 cm^{-1} were selected for data analysis. An average of 27 (461/17) spectra were obtained for each individual with a 3-min integration time per spectrum. Each spectrum was obtained with excitation power ~300 mW 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. For one subject, a mean absolute error (MAE) of 7.8% (RMSECV ~ 0.7 mM) and an R^2 of 0.83 was obtained. When data from 9 volunteers were combined, the MAE was 12.8% with R^2 ~0.7, while combining all 17 volunteers gave MAE ~16.9% (RMSECV ~ 1.5 mM). The number of PLS factors compared to the number of spectra is in danger of overfitting in an individual calibration. However, the grouping schemes

involving 9 (244 spectra) and 17 (461 spectra) volunteers utilized 17 and 21 factors, respectively, which is more acceptable. 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.

This study was 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. 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.

8. Toward prospective application

The results from the *in vitro* and *in vivo* studies reviewed in the previous sections are very encouraging. They demonstrate the feasibility of building glucose-specific *in vivo* multivariate calibration models based on Raman spectroscopy. To bring this technique to the next level, prospective application of a calibration algorithm on independent data with clinically acceptable detection results needs to be demonstrated. From our perspective, this objective requires advances to be made in extracting glucose information without spurious correlations to other system components and correcting for variations in subject subject tissue morphology and color. We have developed new tools to address these issues. Specifically, a novel multivariate calibration technique with higher analyte specificity that is more robust against interferent co-variation or chance correlation was

developed. This technique, constrained regularization, is described in section 8.1. Also, a new correction method to compensate for turbidity-induced sampling volume variations across sites and individuals was developed. This method, intrinsic Raman spectroscopy, is introduced in section 8.2. Additionally, other considerations for successful *in vivo* studies such as reference concentration accuracy, optimal collection site determination, etc., will be discussed in the context of future directions.

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 owing to the analyte or to systematic or spurious effects. One way to effectively boost the model specificity is through incorporation of 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.

8.1.1 Hybrid linear analysis (HLA)

Hybrid linear analysis was developed by Berger *et al.*³⁸ First, analyte spectral contributions are 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 with significant principal components extracted. These principal components are subsequently used as basis spectra to perform an orthogonalization process on the pure analyte spectrum. The 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 the incorporation of the pure glucose spectrum in the algorithm development. However, because HLA relies on the subtraction of the analyte spectrum from the calibration data, it is highly sensitive to the accuracy of the spectral shape and its intensity. For complex turbid samples in which absorption and scattering are likely to alter the analyte spectral features in unknown ways, 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.

8.1.2 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:

$$\mathbf{c} = \mathbf{S}^T \mathbf{b}. \quad (4)$$

The goal is to invert Eq. (4) and obtain a solution for \mathbf{b} . Factor-based methods such as principal component regression (PCR) and partial least squares (PLS) summarize the calibration data, $[\mathbf{S}, \mathbf{c}]$, using a few principal components or loading vectors. Whereas constrain regularization (CR) seeks a balance between model approximation error and noise propagation error by minimizing the cost function, Φ :³⁹

$$\Phi(\Lambda, \mathbf{b}_0) = \|\mathbf{S}^T \mathbf{b} - \mathbf{c}\|^2 + \Lambda \|\mathbf{b} - \mathbf{b}_0\|^2, \quad (5)$$

with $\|\mathbf{a}\|$ the Euclidean norm (*i.e.*, magnitude) of \mathbf{a} , and \mathbf{b}_0 a spectral constraint that introduces prior information about \mathbf{b} . The first term of Φ 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 \mathbf{b} in the absence of noise and interferences. Another choice is the net analyte signal⁴⁰ calculated using all of the known pure analyte spectra. Such flexibility in the selection of \mathbf{b}_0 is owing to the manner in which the constraint is incorporated into the calibration algorithm. For CR, the spectral constraint is included in a nonlinear fashion through minimization of 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.

In numerical simulations and experiments with tissue phantoms, we found that with CR the RMSEP is lower than methods without prior information, such as PLS, and 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

An important lesson learned from the study is that there is a trade-off between maximizing prior information utilization and robustness concerning the accuracy of such information. Multivariate calibration methods range from explicit methods with maximum use of prior information (*e.g.* OLS, least robust when accurate model is not obtainable), 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 is prone to be misled by spurious correlations). We believe CR achieves the optimal balance between these ideals in practical situations.

8.2 Sampling volume correction using intrinsic Raman spectroscopy

Sample variability is a critical issue in prospective application. For optical technologies, variations in tissue optical properties such as absorption and scattering coefficients can create distortions in measured spectra. This section provides a brief overview of techniques to correct turbidity-induced spectral and intensity distortions in fluorescence and Raman spectroscopy, respectively. In particular, photon migration theory is presented as an analytical tool to model diffuse reflectance, fluorescence and Raman scattering arising from turbid biological samples. Monte Carlo simulation is introduced 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 intrinsic Raman spectroscopy are presented.

8.2.1 Optical properties biological tissue

Light propagation in turbid media such as biological tissue is governed by elastic scattering and absorption of the media. Elastic scattering is a phenomenon in which the direction of the photon is changed but not its energy, usually owing to discontinuities in material properties (*e.g.* refractive index) in the media, and absorption is the conversion of light energy into another form of energy (usually thermal energy). Most analytical and numerical models employ macroscopic optical properties, including the absorption coefficient, μ_a (cm^{-1}), the scattering coefficient, μ_s (cm^{-1}), the single scattering angle θ , and the elastic scattering anisotropy, $g = \langle \cos\theta \rangle$, average cosine of the single scattering

angle. The absorption and scattering coefficients are the probability of a photon being absorbed or scattered per unit path length. The sum of μ_a and μ_s is called the total attenuation coefficient, μ_t , with its inverse defined as the mean free depth. The phase function is a probability density function of the scattering deflection angle, describing the probability of a scattering angle at which single scattering event occurs. For example, the Henyey-Greestein phase function⁴¹ is often used to approximate tissue scattering. In general, these optical properties are wavelength dependent.

Optical properties of biological tissue are known to be affected by physiological conditions, tissue morphology, and laser irradiation. Different levels of hematocrit (red blood cells) in whole blood cause different absorption (hemoglobin) and scattering (red cells) properties. Similarly, different skin layer thickness, morphology, and melanin content cause optical turbidity to vary. Such turbidity variations exist across different tissue sites or individuals, and are generally slowly-varying in time. On the other hand, laser irradiation can cause shorter time scale temporal variations in turbidity, typically as a result of heating.⁴²

A limiting factor in noninvasive optical technology is variations in the optical properties of samples under investigation that result in spectral distortions⁴³⁻⁴⁷ and sampling volume (effective optical pathlength) variability.⁴⁸⁻⁵³ These variations will impact a noninvasive optical technique not only in interpretation of spectral features, but also in the construction and application of a multivariate calibration model if such variations are not accounted for. As a result, correction methods need to be developed and applied before further quantitative analysis. For Raman spectroscopy, relatively few correction methods

appear in the literature, and most of them are not readily applicable to biological tissue.⁵⁴⁻

58

In fluorescence spectroscopy, however, diffuse reflectance correction of spectral distortions in biological media has been studied extensively. Analytical models based on photon migration theory⁴³, diffusion theory^{45,59,60}, as well as empirical models⁶¹, have been reported to obtain “intrinsic fluorescence.” In the following, we will review a particular correction method based on photon migration theory for fluorescence spectroscopy, and introduce its Raman counterpart.

8.2.2 Corrections based on photon migration theory

Light propagation in turbid media can be described by the radiative transfer equation.⁶²

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 experience multiple scattering events.⁶² Another very useful approximation is photon migration theory, developed by Wu *et al.*^{43, 44} This method employs probabilistic concepts to describe the scattering of light and to set up a framework that allows the calculation of the diffuse reflectance from semi-infinite turbid media. The total diffuse reflectance from a semi-infinite medium can be written as:

$$R_d \approx \sum_{n=1}^{\infty} f_n(g) * a^n, \quad (6)$$

with $f_n(g)$ the photon escape probability distribution, n the number of scattering events before escaping, g the scattering anisotropy, and a the albedo ($\mu_s/(\mu_s + \mu_a)$). Two fundamental assumptions are made: the photon escape probability distribution of a semi-infinite medium only depends on the number of scattering events and anisotropy; the

lineshape of the escape probability distribution can be approximated by exponential function, *i.e.*, $f_n(g) = k(g)e^{-k(g)n}$. These assumptions are validated by Monte Carlo modeling.

In the same paper, Wu *et al.* derived an analytical equation relating measured fluorescence (F) to the intrinsic fluorescence (IF), defined as the fluorescence as measured from a optically-thin slice of tissue, through diffuse reflectance (R):^{43, 44}

$$IF \approx \frac{1}{(1 - a_x)} F \frac{a_x - a_m}{R_x - R_m}. \quad (7)$$

with IF the intrinsic fluorescence, a the albedo, and R the diffuse reflectance. Subscripts $_x$ and $_m$ denote quantities at the excitation and emission wavelengths, respectively.

This equation and its variants have been employed to recover turbidity-free fluorescence spectra from various types of tissue. The correction facilitates interpretation of underlying fluorophores and consequently improves the accuracy of disease diagnosis.^{43, 44, 47}

The same general principle that applies for intrinsic fluorescence should hold true for Raman spectroscopy as well. Unlike in fluorescence spectroscopy, spectral distortion owing to prominent absorbers is less of an issue in the NIR wavelength range. However, for quantitative analysis the turbidity-induced sampling volume variations become very significant and usually dominate over spectral distortions.

An equation analog to Eq. (7) can be derived for the intrinsic Raman signal (IR) under semi-infinite conditions (sample extends into a half plane and all unabsorbed photons eventually exit the air/sample interface):

$$\text{IR} \approx \mu_{t,x} \text{Ram} \frac{a_x - a_R}{R_x - R_R}. \quad (8)$$

Because most Raman instruments rely on a notch filter to prevent CCD saturation from elastically scattered light, diffuse reflectance at the excitation wavelength is not directly available. Monte Carlo simulations and experimental results show that the intrinsic Raman signal for arbitrary samples, as well as collection geometries, can be more conveniently described by:

$$\text{IR} \approx \mu_{t,x} \frac{\text{Ram}}{(a + bR_m^c)}. \quad (9)$$

Parameters in Eq. (9) can be experimentally calibrated and employed to obtain the intrinsic Raman signal.

8.2.3 Monte Carlo method

Monte Carlo simulation is a statistical tool based on macroscopic optical properties that are assumed to extend uniformly over small units of tissue volume (*i.e.*, a voxel). A pre-defined grid is employed to simulate photon-tissue interaction sites. The mean free path of the photon-tissue interaction sites typically ranges from 10-1000 μm . This method does not consider the details of energy distribution within voxels. Photons are treated as classical particles, and the wave features are neglected.^{62, 63} Since its early introduction as a tool to simulate photon elastic scattering, capabilities such as polarization,^{64, 65} temporal resolution,⁶⁶ fluorescence,⁶⁷ and Raman scattering¹⁰ have been developed. Details of the Monte Carlo simulation for diffuse reflectance (the core program) are well documented in the literature.⁶³

8.2.4 Intrinsic Raman spectroscopy (IRS)

To test Eq. (8), the product $(R_m \times \mu_t)$ is plotted versus the ratio $(R_x - R_m)/(a_x - a_m)$ in Figure 8 using results from Monte Carlo simulations. The intrinsic Raman signal can be obtained from the slope of the linear fit. Note that Eq. (8) is only legitimate when the semi-infinite condition holds, but expression Eq.(9) should be valid for any sample geometry. To test Eq. (9), the product $(R_m \times \mu_t)$ is plotted versus R_m in Figure 9 using results from Monte Carlo simulations. The fit to this curve is the intrinsic Raman signal and can be used to correct for sampling volume variations. It can be seen that this expression fits less well in the presence of high absorption (lower Raman and reflectance). However, such high absorption cases in general are rare in biological tissue in the NIR spectral region.

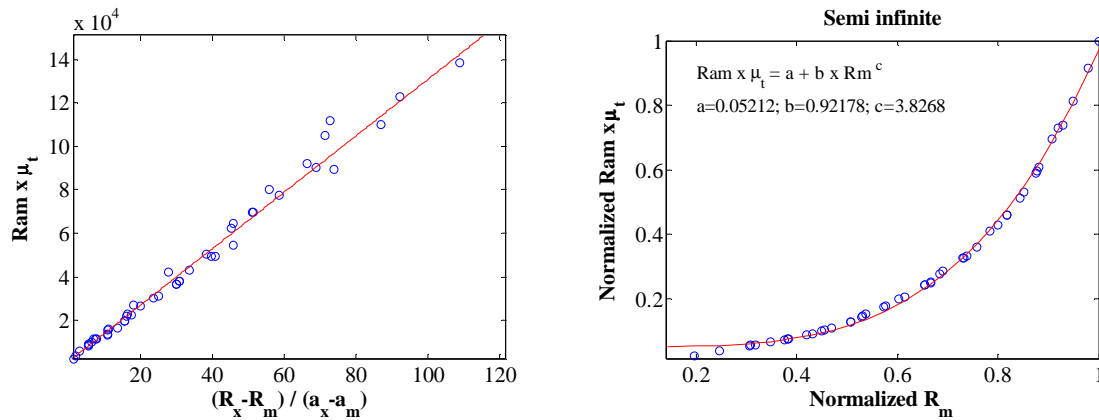


Figure 8 $(R_m \times \mu_t)$ versus $(R_x - R_m)/(a_x - a_m)$. The slope is the intrinsic Raman signal. Figure 9 $(R_m \times \mu_t)$ versus R_m . The fit to the curve can be used to correct for sampling volume variation.

To apply IRS, one needs to know μ_t of the samples. Extraction of optical properties has been studied by many researchers.⁶⁸⁻⁷² The majority of methods are based on diffusion theory or variants of it. Our laboratory extracts optical properties from biological tissue

routinely in other wavelength ranges and a similar method could be employed for this purpose.⁷²

8.3 Other considerations and future directions

The results presented above are promising. However, for non-invasive Raman spectroscopy to be applied prospectively with clinically acceptable accuracy, several additional modifications/improvements/advances/ need to be implemented. We address these below.

Accurate reference concentration measurements

An additional factor that greatly affects the performance of the calibration algorithm is the accuracy of the reference measurements. In spectroscopic techniques such as Raman, 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 interstitial glucose lags the plasma glucose concentration from 5 to 30 minutes in humans.⁷³ As a result, using plasma glucose as the reference concentration may introduce errors. Methods of extracting interstitial fluid for glucose reference measurements should be explored.

Background signal and its variations over time

As mentioned earlier, the intense background, typically described as fluorescence, can limit the detection accuracy in three aspects: the noise associated with the background decreases the SNR; the changes in its spectral shape over time confounds the calibration algorithm; and its intensity variations over time introduces non analyte-specific correlation into the calibration model. Unfortunately, the background-associated noise

can not be removed by background removal techniques. Further, it has been found that removing the background using polynomial fitting does not improve calibration results, potentially owing to non-analyte-specific artifacts.⁸ Thus, methods to reduce the background signal at its origin should be explored. One approach may be using photobleaching combined with intentional motion by, for example, scanning the illumination spot around an area slightly larger than the spot itself. With such a scheme, the apparent background can be lower to start with, and the photobleaching can be reduced.

Optimal probing depth through accurate sample positioning

The probing depth and sample positioning are critical for optimal collection of glucose-specific Raman scattered photons and calibration transfer. In experiments, the optimal probing depth can be estimated from extracted optical properties, and therefore the correct distance between the sample-and the collection optic can be determined for each measurement site. To address this, a fundamental study of morphological and layer structures at the probing site should be carried out with a computer-controlled 3-axis precision stage, as has been done on particular parts of skin.⁹ Because most Raman scatterers have specific spatial distribution in skin, such as keratin in the epidermis, collagen in the dermis, *etc.*, a two-layer model can be developed and utilized. Given such distinctive spatial distributions between keratin and collagen, we can obtain information about the probing depth and even layer thickness by comparing the relative magnitude of keratin and collagen Raman signals. By knowing the exact sampling volume and its coverage of various skin morphological structures, we can estimate how much of the

glucose-containing region (*dermis* in the two-layer model) is sampled. This information can effectively lead to better reference concentrations, improving the calibration accuracy.

Motion artifacts and skin heterogeneity

A key component to obtaining accurate and robust calibrations is the sample interface. The sample interface should ideally limit motion while maintaining a constant pressure and temperature. One approach to combat inadvertent motion artifacts is to intentionally build motion into the calibration model. This can be achieved by scanning the laser spot within a larger area.

Optimal data collection site

Individual calibration models based on cross validation can be established for several candidate sites such as forearm, fingernail, etc, and the results can be compared. The minimum detection error analysis can also be employed to evaluate different sites.

9. Conclusion

Quantitative Raman spectroscopy is a promising technique for noninvasive glucose sensing. From its early development with *in vitro* studies by several groups, *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. For Raman spectroscopy to be a viable clinical technique, successful prospective studies must be carried out. From our perspective, breakthroughs have to be made in the following directions: enhancing glucose specificity, correcting for diversity across individuals, accurate reference

concentration measurements, reducing the fluorescence background, sample positioning and interface, and optimal site determination.

In this chapter we presented our research efforts addressing the first two issues with constrained regularization and intrinsic Raman spectroscopy, respectively. These techniques will play a critical role in prospective studies involving multiple sites/subjects/days. We are currently planning for a multiple-subject and multiple-day *in vivo* study, first on dogs and then on humans. We believe these new developments together with a robust sample interface will enable us to demonstrate prospective applicability.

References

- (1) Raman, C. V.; Krishnan, K. S. *Nature* **1928**, *121*, 501-502.
- (2) McCreery, R. L. *Raman spectroscopy for chemical analysis*; John Wiley & Sons: New York, 2000.
- (3) Qu, J. N. Y.; Wilson, B. C.; Suria, D. *Applied Optics* **1999**, *38*, 5491-5498.
- (4) Baraga, J. J.; Feld, M. S.; Rava, R. P. *Applied Spectroscopy* **1992**, *46*, 187-190.
- (5) Gornushkin, I. B.; Eagan, P. E.; Novikov, A. B.; Smith, B. W.; Winefordner, J. D. *Applied Spectroscopy* **2003**, *57*, 197-207.
- (6) Lieber, C. A.; Mahadevan-Jansen, A. *Applied Spectroscopy* **2003**, *57*, 1363-1367.
- (7) Vickers, T. J.; Wambles, R. E.; Mann, C. K. *Applied Spectroscopy* **2001**, *55*, 389-393.
- (8) Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W. C.; Sasic, S.; Horowitz, G. L.; Feld, M. S. *Journal of Biomedical Optics* **2005**, *10*.
- (9) Caspers, P. J.; Lucassen, G. W.; Puppels, G. J. *Biophysical Journal* **2003**, *85*, 572-580.
- (10) Koo, T.-W. *Measurement of blood analytes in turbid biological tissue using near-infrared Raman spectroscopy*; Massachusetts Institute of Technology: Cambridge, 2001.
- (11) Scepanovic, O.; Bechtel, K. L.; Haka, A. S.; Shih, W. C.; Koo, T.-W.; Feld, M. S., *under review*.
- (12) Roe, J. N.; Smoller, B. R. *Critical Reviews in Therapeutic Drug Carrier Systems* **1998**, *15*, 199-241.
- (13) Guyton, A. C.; Hall, J. E. *Human physiology and mechanisms of disease*, 6th ed.; Saunders: Philadelphia, 1997.
- (14) Martens, H.; Naes, T. *Multivariate Calibration*; John Wiley & Sons: New York, 1989.
- (15) Geladi, P.; Kowalski, B. R. *Analytica Chimica Acta* **1986**, *185*, 1-17.
- (16) Kowalski, B. R.; Lorber, A. *Abstracts of Papers of the American Chemical Society* **1988**, *196*, 100-Anyl.
- (17) Hirschfeld, T.; Chase, B. *Applied Spectroscopy* **1986**, *40*, 133-137.
- (18) Wang, Y.; McCreery, R. L. *Analytical Chemistry* **1989**, *61*, 2647-2651.
- (19) Qu, J. Y.; Shao, L. *Review of Scientific Instruments* **2001**, *72*, 2717-2723.
- (20) Lambert, J. L.; Morookian, J. M.; Sirk, S. J.; Borchert, M. S. *Journal of Raman Spectroscopy* **2002**, *33*, 524-529.
- (21) Lambert, J. L.; Pelletier, C. C.; Borchert, M. *Journal of Biomedical Optics* **2005**, *10*.

- (22) Enejder, A. M. K.; Koo, T. W.; Oh, J.; Hunter, M.; Sasic, S.; Feld, M. S.; Horowitz, G. L. *Optics Letters* **2002**, *27*, 2004-2006.
- (23) Berger, A. J.; Koo, T. W.; Itzkan, I.; Horowitz, G.; Feld, M. S. *Applied Optics* **1999**, *38*, 2916-2926.
- (24) Chaiken, J.; Finney, W.; Knudson, P. E.; Weinstock, R. S.; Khan, M.; Bussjager, R. J.; Hagrman, D.; Hagrman, P.; Zhao, Y. W.; Peterson, C. M.; Peterson, K. *Journal of Biomedical Optics* **2005**, *10*.
- (25) Motz, J. T.; Hunter, M.; Galindo, L. H.; Gardecki, J. A.; Kramer, J. R.; Dasari, R. R.; Feld, M. S. *Applied Optics* **2004**, *43*, 542-554.
- (26) Huang, Z. W.; Zeng, H. S.; Hamzavi, I.; McLean, D. I.; Lui, H. *Optics Letters* **2001**, *26*, 1782-1784.
- (27) Shih, W. C.; Bechtel, K. L.; Feld, M. S. *Analytical Chemistry* **2007**, *79*, 234-239.
- (28) Zhao, J. *Applied Spectroscopy* **2003**, *57*, 1368-1375.
- (29) Chrisp, M. P., 1999.
- (30) Pelletier, I.; Pellerin, C.; Chase, D. B.; Rabolt, J. F. *Applied Spectroscopy* **2005**, *59*, 156-163.
- (31) Pelletier, C. C.; Lambert, J. L.; Borchert, M. *Applied Spectroscopy* **2005**, *59*, 1024-1031.
- (32) Berger, A. J.; Wang, Y.; Sammeth, D. M.; Itzkan, I.; Kneipp, K.; Feld, M. S. *Applied Spectroscopy* **1995**, *49*, 1164-1169.
- (33) Berger, A. J.; Wang, Y.; Feld, M. S. *Applied Optics* **1996**, *35*, 209-212.
- (34) Berger, A. J.; Itzkan, I.; Feld, M. S. *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy* **1997**, *53*, 287-292.
- (35) Rohleder, D.; Kiefer, W.; Petrich, W. *Analyst* **2004**, *129*, 906-911.
- (36) Chaiken, J.; Finney, W. F.; Yang, X.; Knudson, P. E.; Peterson, K. P.; Peterson, C. M.; Weinstock, R. S.; Hagrman, D. *Proceedings of SPIE* **2001**, *4254*, 216-227.
- (37) Chaiken, J.; Peterson, C.: U.S., 2001.
- (38) Berger, A. J.; Koo, T. W.; Itzkan, I.; Feld, M. S. *Analytical Chemistry* **1998**, *70*, 623-627.
- (39) Bertero, M.; Boccacci, P. *Introduction to inverse problems in imaging*; Institute of Physics Pub.: Bristol, UK ; Philadelphia, PA, 1998.
- (40) Lorber, A. *Analytical Chemistry* **1986**, *58*, 1167-1172.
- (41) Henyey, L.; Greenstein, J. *Astrophysical Journal* **1941**, *93*, 70-83.
- (42) Welch, A. J.; van Gemert, M. J. C. *Optical-thermal response of laser-irradiated tissue*; Plenum Press, 1995.
- (43) Wu, J.; Feld, M. S.; Rava, R. P. *Applied Optics* **1993**, *32*, 3585-3595.

- (44) Zhang, Q. G.; Muller, M. G.; Wu, J.; Feld, M. S. *Optics Letters* **2000**, *25*, 1451-1453.
- (45) Zhadin, N. N.; Alfano, R. R. *Journal of Biomedical Optics* **1998**, *3*, 171-186.
- (46) Gardner, C. M.; Jacques, S. L.; Welch, A. J. *Applied Optics* **1996**, *35*, 1780-1792.
- (47) Muller, M. G.; Georgakoudi, I.; Zhang, Q. G.; Wu, J.; Feld, M. S. *Applied Optics* **2001**, *40*, 4633-4646.
- (48) Weersink, R.; Patterson, M. S.; Diamond, K.; Silver, S.; Padgett, N. *Applied Optics* **2001**, *40*, 6389-6395.
- (49) Arnold, M. A.; Small, G. W. *Analytical Chemistry* **2005**, *77*, 5429-5439.
- (50) Cote, G. L.; Lec, R. M.; Pishko, M. V. *Ieee Sensors Journal* **2003**, *3*, 251-266.
- (51) Khalil, O. S. *Clinical Chemistry* **1999**, *45*, 165-177.
- (52) Diamond, K. R.; Farrell, T. J.; Patterson, M. S. *Physics in Medicine and Biology* **2003**, *48*, 4135-4149.
- (53) Pogue, B. W.; Burke, G. *Applied Optics* **1998**, *37*, 7429-7436.
- (54) Nijhuis, T. A.; Tinnemans, S. J.; Visser, T.; Weckhuysen, B. M. *Physical Chemistry Chemical Physics* **2003**, *5*, 4361-4365.
- (55) Waters, D. N. *Spectrochimica Acta Part a-Molecular and Biomolecular Spectroscopy* **1994**, *50*, 1833-1840.
- (56) Aarnoutse, P. J.; Westerhuis, J. A. *Analytical Chemistry* **2005**, *77*, 1228-1236.
- (57) Kuba, S.; Knozinger, H. *Journal of Raman Spectroscopy* **2002**, *33*, 325-332.
- (58) Tinnemans, S. J.; Kox, M. H. F.; Nijhuis, T. A.; Visser, T.; Weckhuysen, B. M. *Physical Chemistry Chemical Physics* **2005**, *7*, 211-216.
- (59) Patterson, M. S.; Pogue, B. W. *Applied Optics* **1994**, *33*, 1963-1974.
- (60) Finlay, J. C.; Foster, T. H. *Applied Optics* **2005**, *44*, 1917-1933.
- (61) Finlay, J. C.; Conover, D. L.; Hull, E. L.; Foster, T. H. *Photochemistry and Photobiology* **2001**, *73*, 54-63.
- (62) Ishimaru, A. *Wave propagation and scattering in random media*; Academic Press: New York, 1978.
- (63) Wang, L. H.; Jacques, S. L.; Zheng, L. Q. *Computer Methods and Programs in Biomedicine* **1995**, *47*, 131-146.
- (64) Ramella-Roman, J. C.; Prahl, S. A.; Jacques, S. L. *Optics Express* **2005**, *13*, 4420-4438.
- (65) Ramella-Roman, J. C.; Prahl, S. A.; Jacques, S. L. *Optics Express* **2005**, *13*, 10392-10405.
- (66) Hielscher, A. H.; Liu, H. L.; Chance, B.; Tittel, F. K.; Jacques, S. L. *Applied Optics* **1996**, *35*, 719-728.

- (67) Swartling, J.; Pifferi, A.; Enejder, A. M. K.; Andersson-Engels, S. *Journal of the Optical Society of America a-Optics Image Science and Vision* **2003**, *20*, 714-727.
- (68) Farrell, T. J.; Patterson, M. S.; Wilson, B. *Medical Physics* **1992**, *19*, 879-888.
- (69) Dam, J. S.; Pedersen, C. B.; Dalgaard, T.; Fabricius, P. E.; Aruna, P.; Andersson-Engels, S. *Applied Optics* **2001**, *40*, 1155-1164.
- (70) Doornbos, R. M. P.; Lang, R.; Aalders, M. C.; Cross, F. W.; Sterenborg, H. J. C. M. *Physics in Medicine and Biology* **1999**, *44*, 967-981.
- (71) Nichols, M. G.; Hull, E. L.; Foster, T. H. *Applied Optics* **1997**, *36*, 93-104.
- (72) Zonios, G.; Perelman, L. T.; Backman, V. M.; Manoharan, R.; Fitzmaurice, M.; Van Dam, J.; Feld, M. S. *Applied Optics* **1999**, *38*, 6628-6637.
- (73) Boyne, M. S.; Silver, D. M.; Kaplan, J.; Saudek, C. D. *Diabetes* **2000**, *49*, A99-a99.