Real-time Raman system for in vivo disease diagnosis

Jason T. Motz
Saumil J. Gandhi
Obrad R. Scepanovic
Abigail S. Haka
John R. Kramer
Ramachandra R. Dasari
Michael S. Feld

Massachusetts Institute of Technology
Cambridge, Massachusetts 02138

1 Introduction

Raman spectroscopy is a promising means of studying biological tissue.1-5 The technique relies on extracting chemical information from a sample based on its molecule-specific Raman emission spectrum, which results from molecular vibrations excited by incident monochromatic light. Up to now, most research in this field has studied excised and cadaveric tissue samples.3-6 These reports include the examination of atherosclerosis,3,4 the measurement of blood analytes,7 and detection of dysplasia and cancer in various tissues including breast,8-10 cervix,1 and prostate.11 Clinical applications have been limited in scope, primarily due to the lack of optical fiber probes capable of effectively acquiring Raman spectra from tissue within the body. As a result, the majority of in vivo Raman spectroscopy research has been confined to easily accessible organs and tissues such as skin12-14 and the cervix.15

In order to monitor tissue sites within the body, flexible, small diameter optical fibers must be used to deliver light to the tissue and collect the returning light. Because the probability of Raman scattering is so small, unique challenges must be considered in the design of Raman probes. In particular, two problems must be addressed. First, intense luminescence background generated by the strong monochromatic excitation beam in the delivery and collection optical fibers can mask the signal of interest and generate significant levels of shot noise that compromises the tissue Raman signal. Second, the probe must be capable of efficient, high-throughput collection, an essential requirement for rapid data acquisition.

Previous Raman studies examining internal body tissues with flexible optical fiber probes have required long collection times, in the range of 5-30 s, which are impractical for many clinical applications.16-18 Recent advances in probe design in our laboratory have resulted in a flexible Raman probe capable of collecting spectra with large signal-to-noise ratio in 1 s or less.19 In vivo Raman spectroscopy of human arteries20 and breast tissues has been demonstrated.

Practical implementation of Raman spectroscopy, however, requires an integrated instrument that can provide real-time spectral analysis and diagnostic information to the clinician. A previously described Raman system served as proof of principle for these concepts.21 Using a commercially available probe to take spectra from a human finger, arm, nail, tooth, and tongue with acquisition times of 5 s, custom-built software was able to provide accurate classification of the different tissue types within 1 s after acquisition.

In this paper we present a real-time Raman instrument with which spectra are acquired in 1 s, while analysis and diagnostic information are provided within an additional second. The instrument is designed to work in conjunction with the newly developed optical fiber Raman probes and its performance is illustrated with data from recent clinical studies of atherosclerosis and breast cancer. The ability to rapidly collect Raman diagnostic information, such as morphological and chemical composition,3,4,22,23 and provide immediate feedback to the physician during a clinical procedure has the potential to greatly improve patient care by facilitating real-time disease diagnosis.

Address all correspondence to Jason T. Motz, Wellman Center for Photomedicine, Massachusetts General Hospital, Harvard Medical School, 50 Blossom St., BARK704, Boston, MA 02114. Tel: 617-726-3020; Fax: 617-726-4103. E-mail: jmotz@partners.org

Abstract. Raman spectroscopy has been well established as a powerful in vitro method for studying biological tissue and diagnosing disease. The recent development of efficient, high-throughput, low-background optical fiber Raman probes provides, for the first time, the opportunity to obtain real-time performance in the clinic. We present an instrument for in vivo tissue analysis which is capable of collecting and processing Raman spectra in less than 2 s. This is the first demonstration that data acquisition, analysis, and diagnostics can be performed in clinically relevant times. The instrument is designed to work with the new Raman probes and includes custom written LabVIEW and Matlab programs to provide accurate spectral calibration, analysis, and diagnosis along with important safety features related to laser exposure. The real-time capabilities of the system were demonstrated in vivo during femoral bypass and breast lumpectomy surgeries. Such a system will greatly facilitate the adoption of Raman spectroscopy into clinical research and practice. © 2005 Society of Photo-Optical Instrumentation Engineers. [DOI: 10.1117/1.1920247]

Keywords: in vivo Raman spectroscopy; real-time diagnosis; atherosclerosis; breast cancer; Raman probes.

Paper SS04173RR received Aug. 31, 2004; revised manuscript received Mar. 8, 2005; accepted for publication Mar. 10, 2005; published online May 24, 2005.

1083-3668/2005/$22.00 © 2005 SPIE

Downloaded From: https://www.spiedigitallibrary.org/journals/Journal-of-Biomedical-Optics on 05 Apr 2019
Terms of Use: https://www.spiedigitallibrary.org/terms-of-use
management, improving diagnostic accuracy, and allowing early disease diagnosis and long-term monitoring of therapy.

2 Instrumentation

The excitation source and detection unit of the clinical Raman system used in this study (Fig. 1) is similar to that described by Brennan et al. and Shim and Wilson. Light from an 830 nm InGaAs diode laser (Process Instruments, Salt Lake City, UT) is collimated by a pair of cylindrical lenses and passed through a holographic bandpass filter centered at 830 nm (Kaiser Optical Systems, Ann Arbor, MI). The light is then coupled into the 200 μm core diameter excitation fiber of the Raman probe using a 5x microscope objective (numerical aperture (NA)=0.1, Newport Corporation, Irvine, CA). Illumination of the sample is gated by a high-speed, 6 mm aperture, computer-controlled shutter (LS6ZM2, Vincent Associates, Rochester, NY). The proximal end of the excitation fiber is terminated with an FC connector to provide day-to-day reproducibility of alignment.

A bifurcated optical fiber catheter is used to transport the excitation and Raman scattered light to and from the tissue sample at the distal end of the catheter. Fifteen collection fibers, each with a 200 μm core diameter, surround a single excitation fiber at the distal end of the probe. A short pass filter is placed at the tip of the excitation fiber such that the fiber background is rejected and only 830 nm excitation laser light reaches the tissue. In addition, a long pass filter is placed...
at the tips of the collection fibers in order to minimize the Rayleigh scattered excitation light entering the collection fibers, thereby reducing the generation of additional fiber background. The distal end of the probe is terminated with a sapphire ball lens to enhance signal collection and provide a roughly collimated 200 μm diameter illumination spot. For 100 mW of excitation power, the resultant irradiance is 318 W/cm². Similar and larger power densities have been used in prior studies and shown to not cause a significant increase in temperature or any damage to arterial tissue.26,27

At the proximal end of the probe, the 15 collection fibers are aligned in a linear array, effectively serving as an entrance slit to the spectrograph, resulting in a 9.4 cm⁻¹ spectral resolution. The fibers are encased in a modified BNC connector to ensure highly reproducible alignment with the connector plate of the spectrograph. Any spectral changes incurred by disconnecting and reconnecting the linear array are below the system’s spectral resolution. The numerical aperture of the collection fibers is f/#-matched to the Holospec f/1.8 spectrograph (Kaiser Optical Systems) to conserve throughput. The spectrograph contains an 830 nm notch filter that further suppresses the elastically scattered Rayleigh light and a holographic grating to disperse the Raman scattered light onto a back-illuminated, deep-depletion charge coupled device (CCD) detector with a 1340×400 array of pixels (Spec10:400BR, Princeton Instruments, Trenton, NJ). The CCD detector has a 16 bit dynamic range and was liquid nitrogen cooled to −110 °C. Spectra are obtained by vertically binning the signal from the 15 collection fibers.

3 Real-Time Software

The system utilizes LabVIEW software (v 6.1, National Instruments, Austin, TX) as the primary platform to interface with various hardware components of the system and control data flow, allowing automation of more complicated calibration and analysis routines. These subroutines are accomplished via a direct interface with Matlab (v.6.5.0, The Mathworks, Natick, MA) and provide relevant diagnostic parameters, such as percentage content of calcification and total cholesterol in the case of atherosclerosis, or collagen and fat in the case of breast cancer.3,4,22,23 The modular design of the real-time software allows for easy adaptation of the instrument for different applications.

Fig. 3 *In vitro* Raman spectrum of a non-atherosclerotic femoral artery (a) unprocessed and indicating the biological Raman ROI, (b) after spectral response correction and fiber background removal, and (c) following tissue luminescence removal.
3.1 System Calibration

In order to provide real-time data analysis and diagnosis, spectra for calibration and background subtraction must be acquired prior to data collection. In this study 4-acetamidophenol (Tylenol) was used to calibrate the Raman shift axis. The remaining fiber background signal from the optical fiber probe is characterized by acquisition of a spectrum from roughened aluminum. The spectral response of the collection system is obtained by recording the spectrum of a calibrated tungsten white-light source, diffusely scattered by a reflectance standard (BaSO₄). All calibration spectra are acquired with the same probe to be used for the subsequent procedure and loaded into the system prior to in vivo data collection. Alignment and calibration data can all be obtained within a couple of minutes. Following collection of the calibration spectra, the Raman probe is submitted for sterilization. The system has been designed so that it is stable from day to day, thus disconnecting and reconnecting the probe to the system does not change the calibration parameters.

3.2 Data Acquisition and Analysis

Figure 2 shows the flow of data between different modules of the real-time software and the clinical Raman system. LabVIEW control of the CCD detector is accomplished through a data acquisition (DAQ) card (PCI-6035E, National Instruments, Austin, TX) with drivers written by R²-Software (Lawrenceville, NJ). The right-hand side of Fig. 2 depicts the LabVIEW-Matlab interface for real-time data analysis and disease diagnosis.

A sequence of steps is executed before a Raman diagnosis is extracted from the raw spectrum acquired on the CCD. First, the data are corrected for the system spectral response by dividing them by the white-light spectrum. The resultant data are then normalized to maximum peak height and truncated to the biological Raman region of interest (ROI) appropriate to the models used in this study (686–1788 cm⁻¹). The next step is to remove the fiber background from the tissue data by subtracting the spectrum acquired from roughened aluminum. The relationship between the intensity of spectra from the aluminum and the tissue spectrum is dependent upon the tissue’s optical properties, which are unknown a priori. Therefore, we iteratively subtract the same aluminum spectrum scaled by a range of different intensities to determine the optimal ratio for background removal. The spectrum that results in the lowest standard deviation of the residual between the data and the model fit is used for analysis. After fiber background removal, the remaining broadband tissue luminescence is removed by subtracting a sixth order polynomial that is fit to the data.²⁴ This final correction step results in extraction of the Raman spectrum of the examined site. The resultant spectrum is calibrated to a standard set of wavenumber bins. This spectrum is analyzed with the appropriate disease model by using ordinary least squares (OLS), which provides the relative fit coefficients of the various chemical/morphological components of tissue. Finally, the corresponding classification algorithm is invoked to provide a clinical diagnosis for the region of tissue examined (see Diagnostic Algorithms, below).

The outlined sequence of calibration and data preprocessing steps is illustrated in Fig. 3 with an in vitro spectrum taken from normal femoral artery. The spectrum was collected in four accumulations of 0.25 s for a total collection time of 1 s with 100 mW excitation laser power incident on the sample. As seen in Fig. 3(a), the raw tissue spectrum consists of small tissue Raman features, which are barely discernible above the fiber background, and large, broad tissue luminescence despite the excellent filtering employed in the probe. Figure 3(b) shows the tissue data after its intensity is normalized and corrected for spectral response. The primary effect of this step is to correct for collection system efficiency and the monotonically decreasing quantum efficiency of the CCD at higher wavenumbers. Figure 3(b) also shows the tissue spectrum after final removal of the remaining fiber background. Finally, Fig. 3(c) presents the Raman spectrum of this normal femoral artery after the broadband tissue luminescence has been removed. For consistency with previous studies, the intensity of the spectrum is normalized after tissue luminescence removal.

3.3 Diagnostic Algorithms

The real-time software is configured to allow selection of the appropriate spectral model³,²² and tissue-specific diagnostic algorithms.⁴,²³ This currently includes those developed in our laboratory for diagnosis of atherosclerosis and breast cancer.
These models include basis spectra of morphological and chemical components for OLS fitting of the tissue spectra, as well as diagnostic thresholds for disease classification. The user can toggle between the two different models as mandated by the clinical situation. Once the Raman spectrum is extracted, data are fit by OLS using the appropriate user-selected model and the corresponding diagnosis is presented.

### 3.4 Safety

The *in vivo* Raman system must adhere to strict safety guidelines in the clinical environment. Any part of the system that touches the patient directly or indirectly must be kept sterile. Sterilization during this study was performed using cold-gas ethylene oxide. Furthermore, the excitation laser power cannot exceed a predetermined threshold value.

As shown on the left-hand side of Fig. 2, accurate control of the excitation laser power is accomplished with LabVIEW. After sterilization, the probe tip cannot be contaminated by measuring the excitation power with a power meter, therefore we use a sterilized Teflon block to calibrate the excitation laser power. A Teflon spectrum obtained prior to the procedure is used to determine the expected signal (target intensity) for the desired excitation power. During the procedure, spec-

**Table 1** Timing breakdown for data acquisition, analysis, and disease diagnosis necessitated by the clinical Raman system.

<table>
<thead>
<tr>
<th>Process</th>
<th>Time (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Open laser shutter</td>
<td>0.700</td>
</tr>
<tr>
<td>Collection time</td>
<td>1000</td>
</tr>
<tr>
<td>Load acquired spectra in Matlab</td>
<td>80</td>
</tr>
<tr>
<td>Spectral response correction</td>
<td>1</td>
</tr>
<tr>
<td>Fiber background removal</td>
<td>1</td>
</tr>
<tr>
<td>Tissue luminescence removal</td>
<td>200</td>
</tr>
<tr>
<td>Calibration to standard model bins</td>
<td>480</td>
</tr>
<tr>
<td>Ordinary least squares fitting</td>
<td>190</td>
</tr>
<tr>
<td>Diagnostic algorithm and on-screen display</td>
<td>70</td>
</tr>
</tbody>
</table>
tra of an identical sterilized Teflon block are taken with the same probe used for calibration, and their peak intensities are compared to the target intensity. The laser power is automatically adjusted via LabVIEW until the target intensity is obtained, or until a predetermined threshold power is reached (Fig. 4). Laser power is modified remotely with an analog waveform through the DAQ. The alignment procedure takes less than 1 min and typically results in laser powers within 25% of the target.

As a precaution for safety of the operators and patient, the laser beam is gated by a high-speed shutter, also controlled by LabVIEW. The shutter opens automatically just before data acquisition begins and closes immediately after the acquisition is complete. The shutter then remains closed until the system receives a signal through LabVIEW for the next acquisition.

4 System Performance

The tissue Raman spectrum, model fit, residual between the data and model fit, fractional fit coefficients, and diagnosis are plotted on the computer screen (Fig. 5). Using an IBM compatible computer with a clock speed of 2.8 GHz, the entire process of data acquisition, analysis, and diagnosis is accomplished in less than 2 s. The timing breakdown for each step of the process is shown in Table 1.

Limitations to reducing the total time for data acquisition, analysis, and disease diagnosis are mainly imposed by two factors: collection efficiency and data analysis. First, the collection efficiency and throughput of the optical fiber Raman probe determine the time needed for collecting in vivo tissue spectra. Significant improvements in the design of the probe filters and collection optics will likely lead to higher throughput and collection efficiency, allowing in vivo data acquisition in much less than 1 s.

Reduction in time for data analysis is mainly limited by processor speed. As can be seen in Table 1, the most time consuming steps for data analysis are tissue luminescence removal, data calibration, and OLS fitting of the data. The relationship between the intensity of the aluminum spectrum and the tissue spectrum is unknown a priori. Therefore, the same aluminum spectrum scaled with various intensities is needed to determine the optimal ratio for fiber background removal. As a result, OLS fitting has to be iteratively carried out for each tissue site. Development of algorithms that rely on knowledge of a priori probabilities of probing a certain tissue type might further reduce time spent on analyzing the data. In addition, more efficient filtering in future Raman probe designs will further reduce the fiber background and therefore decrease the uncertainty about its magnitude. Calibrating the data to the standard wavenumber bins utilized in the spectral models is the most time consuming step in analysis. It is possible that switching to a more efficient programming platform could provide significant improvements to this performance.

5 Real-Time Diagnosis

We have used the real-time clinical Raman system for in vivo studies to diagnose atherosclerosis and vulnerable plaque,20 and also to provide margin assessment during breast cancer surgery. Data were collected in 1 s with ~100 mW excitation power and representative spectra are shown in Fig. 6. All clinical studies were approved by the MIT Committee on the Use Of Humans as Experimental Subjects and the Institutional Review Board of the cooperating hospitals. Informed consent was obtained from all subjects prior to procedures.

6 Conclusion

A Raman spectroscopy system capable of providing real-time disease diagnosis in a clinical setting has been designed, developed, and tested. The implemented software package utilizes LabVIEW to interface with various components in the system. LabVIEW control of system components such as the laser and laser shutter allows data acquisition in accordance with clinical safety guidelines. In addition, we utilize a LabVIEW-Matlab interface to combine the computational power and versatility of both programs to perform real-time data analysis, while maintaining a customizable, user-friendly
interface. The modular design of the system allows easy incorporation of various spectral models. Furthermore, it provides easily adaptable data analysis routines and diagnostic algorithms.

Acknowledgments
The authors thank Dr. Arnold Miller, Dr. Maryann Fitzmaurice, and Dr. Joseph Crowe for their assistance in clinical procedures and pathological diagnosis. This research was sponsored by the NIH (Grant No. R01-HL-64675) and the National Center for Research Resources program (Grant No. P41-RR-02594). Support was also generously contributed by the Pfizer Corporation and the Lord Foundation of MIT.

References