Quantitative concentration measurements of creatinine dissolved in water and urine using Raman spectroscopy and a liquid core optical fiber

Dahu Qi
University of Rochester
The Institute of Optics
Department of Physics and Astronomy
Rochester, New York 14627

Andrew J. Berger
University of Rochester
The Institute of Optics
Rochester, New York 14627

Abstract. We describe the use of a Teflon®-AF liquid core optical fiber (LCOF) geometry to enhance the collection of Raman scattering from the biochemical creatinine, dissolved in water and in urine. At short integration times, where shot noise is most troublesome, the enhanced signal leads to greater accuracy in estimating the creatinine concentration from the spectrum. At longer integration times, instabilities in the LCOF geometry manifest themselves, and the predictions are the same as or worse than those from standard cuvette-based spectral measurements. Photobleaching of fluorescence from urine is more extensive and more stable in the LCOF as well. Starting from the measured enhancement of a major creatinine Raman band, we calculate the expected ratio of prediction errors obtained using the two geometries, and it agrees closely with the observed ratio. These results indicate that Raman spectroscopy with these Teflon®-AF LCOFs is stable enough for quantitative concentration predictions, accurate to a few percent of the concentration range spanned. [DOI: 10.1117/1.1917842]

Keywords: Raman; spectroscopy; LCOF; waveguide; urine; PLS.

Paper SS04168 received Aug. 27, 2004; revised manuscript received Oct. 29, 2004; accepted for publication Oct. 29, 2004; published online May 11, 2005.

1 Introduction

Raman spectroscopy is a useful method for nondestructive, reagentless measurement of chemical concentrations. In the biomedical arena, Raman spectroscopy has been widely used for quantitative concentration measurement of key chemicals, or ''analytes,'' in solid and liquid biological specimens;[1–9] for a recent review of the field, see Hanlon et al.[10]

With improvements in lasers, filters, spectrographs, and detectors, the main sources of spectral noise in many biological Raman measurements are autofluorescence and other emissions from the sample itself. In such cases, the spectral signal to noise ratio (SNR) can dictate the level of concentration prediction error. Maximizing the SNR therefore becomes essential.

For a given sample, the choice of excitation wavelength determines the intrinsic ratio of background emission to Raman signal. For optimal SNR, this ratio should be as low as possible. In biological applications, where most fluorophores emit strongly in the visible, this often drives the use of near-infrared excitation. As wavelength increases, however, the Raman cross section drops and silicon detectors become less efficient, placing some constraints on the excitation wavelength.

Once the excitation wavelength and the major equipment pieces are chosen, the only ways to improve the SNR are to (a) increase the number of excitation photons delivered to the sample, (b) improve the excitation/collection efficiency, or (c) decrease the fluorescence through photobleaching. Of these, (a) is a brute-force approach that can be accomplished by increasing the laser power and/or the integration time, but both of these variables have practical limits.

For optically clear liquids, a way of addressing both (b) and (c) is by placing the sample in the interior of an optical waveguide, sometimes called a liquid core optical fiber (LCOF).[11] Until recently, such waveguiding was not possible for aqueous samples, due to the lack of suitable materials with refractive indices lower that of water ($n \approx 1.33$). Recently, however, the DuPont company's development of Teflon®-AF, an amorphous fluoropolymer with refractive index $n \approx 1.29$ in the visible and near-infrared, has enabled true waveguiding to be performed in aqueous samples for the first time. Several groups have reported the use of Teflon®-AF for enhanced Raman signals of up to two orders of magnitude over more conventional sample holders, such as 1 cm pathlength cuvettes or similarly sized cylindrical vials.[12–18]

The LCOF greatly increases the pathlength of optical interaction (more total Raman scattering) and also guides (some) Raman scattering from all regions of the LCOF to a common aperture for convenient collection. This improves the SNR by increasing the number of detected photons, just as in approach (a). Additionally, by illuminating the entire sample, more extensive photobleaching can occur, further improving the SNR.[19] These effects should therefore enable more accu-

Address all correspondence to Andrew J. Berger. Tel: 585-273-4724; Fax: 585-244-4936; E-mail: ajberger@optics.rochester.edu

1083-3668/2005/$22.00 © 2005 SPIE

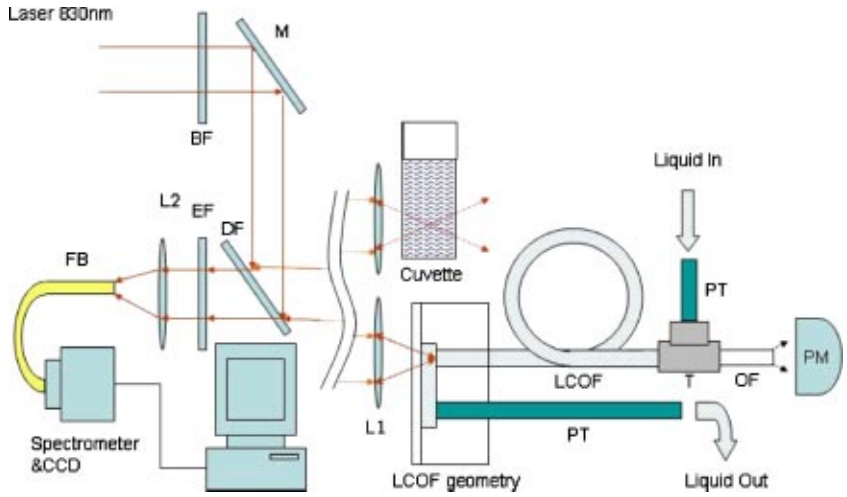

Fig. 1 Block diagram of Raman spectroscopy system, allowing measurements in LCOF or cuvette. Symbols: BF, bandpass filter; *M*, mirror; FB, fiber bundle; *L*1, *L*2, lenses; EF, edge filter; DF, dichroic filter; PT, plastic tubing; *T*, T connection; OF, optical fiber; PM, power meter. See text for details.

rate concentration predictions to be extracted from measured spectra.

Systems with LCOFs are, however, more difficult to align. The launch of light into the LCOF is affected by slight displacements of the tubing during injection of the sample. Even if the tubing does not move, small air bubbles in the liquid scatter light in irreproducible ways from sample to sample. Furthermore, manufacturing imperfections lead to inhomogeneities in the walls of the LCOF, i.e., local regions of higher diffuse scattering loss. Slight changes in the launch of the light can strongly affect how these regions are sampled. Consequently, the enhancement does not remain perfectly stable from one spectrum to the next. For the same set of reasons, the spectral background can change noticeably and abruptly.

Both of these effects (variability in enhancement and in background) work counter to the improvement in SNR and tend to degrade prediction accuracy relative to that achieved using the more stable cuvette geometry. In a given practical case, it is therefore uncertain *a priori* which geometry is preferable.

As mentioned earlier, many reports of enhancements and increased sensitivity have been made, but to our knowledge there have not been investigations of (a) whether a LCOF geometry is stable and reproducible enough to enable quantitative concentration predictions (within a few percent) and (b) the relative concentration prediction performance of LCOF and cuvette geometries when measuring the same samples.

Below, we report on efforts to extract predictions of creatinine concentration, in both water and spiked human urine, from LCOF-based Raman measurements. We also present side-by-side results of LCOF and cuvette-based prediction accuracy as a function of integration time.

2 Methods

2.1 *Optical System*

The Raman spectroscopy system is depicted in Fig. 1 and has been described elsewhere in detail.[20] 830 nm laser light (Process Instruments, Salt Lake City, UT) is delivered to the sample area using 2 m of multimode optical fiber (Thorlabs, Inc., Newton, NJ, 100 μm core diameter). The laser is expanded to a 1 in. diameter beam using a lens ($f=45$ mm, $D=25.4$ mm), filtered by a bandpass filter (Chroma, Burlington, VT) to extinguish Raman signals from the fiber, reflected by a dichroic beam splitter (Chroma, Burlington, VT), and then focused into the liquid sample (either LCOF or cuvette) by lens 1 ($f=35$ mm, $D=25.4$ mm). The LCOF (Biogeneral, Inc.) is a flexible tube of Teflon®-AF (DuPont) with inner diameter 600 μm, outer diameter 800 μm, length 20 cm, and refractive index approximately 1.29. The spot size at the focus point was measured to be 200 μm (>95% of total power), which is much smaller than the inner diameter of the LCOF tubing, ensuring high coupling efficiency. The half angle of the converging light is 9°, matched to the acceptance angle of the waveguide (9.1° when filled with water). The tubing is held in a homemade base, with the light entering through a quartz window placed in front of the LCOF. This base is magnetically mounted and can be swapped with a second base that holds a standard 1×1×4 cm quartz cuvette, as indicated in the figure, positioned so the laser is focused at the center of the cuvette. Backscattered light from the cuvette or the LCOF is collected and collimated by lens 1, long-passed by the dichroic beam splitter, notch-filtered (Chroma, Burlington, VT), and focused by lens 2 ($f=60$ mm, $D=24.5$ mm) onto a 1 mm diameter fiber bundle (FiberGuide, Stirling, NJ), which consists of 41 multimode fibers with 100/140 μm core/cladding diameter. The spot size was calculated to be 700 μm, which illuminates several of the collection fibers without risk of overfilling the probe's face. The other side of the bundle is arranged into a linear array at the entrance slit of an *f*-matched spectrometer (f/1.8 Kaiser Holospec, Ann Arbor, MI) and an open-electrode charge coupled device (CCD) detector (Andor Technology, South Windsor, CT).

The LCOF was connected to an optical fiber and plastic tubing using a T connection, as shown. The optical fiber transfers the laser light that passes through the LCOF to a power meter, for alignment purposes. Sample was injected into the plastic tubing by a syringe, sent through the LCOF via the T connection, and subsequently brought to a disposal beaker via an exit port.

2.2 *Sample Preparation*

Two sets of samples were made for the experiment. In each case, two chemicals were used: creatinine (Sigma-Aldrich, St. Louis, MO), a key component of urine, as the target analyte, and ethanol as an interferent. These were dissolved in two host liquids, either de-ionized (DI) water or urine, to make stock solutions. Samples were made by mixing different amounts of these solutions and the pure host liquid using a 0.5 mL pipette.

Twenty-one samples were made using DI water as the host. The creatinine concentrations varied from 0 to 150 mg/dL, in increments of 7.5 mg/dL, and the ethanol concentrations varied from 0% to 1% of total volume in increments of 0.05%. The two concentrations were assigned to the samples in an uncorrelated fashion (correlation coefficient -0.039). The range of creatinine concentrations corresponds to physiological variation ranges, and the ethanol concentration range was chosen to give approximately the same level of Raman signal variation as the target analyte.

Twenty-one other samples were made using urine. The urine was collected from a healthy male volunteer and then spiked with various levels of creatinine and ethanol as described earlier. The native creatinine level was not measured and was not relevant to the study performed, because it was constant for all samples and only deviations from the mean were eventually calculated. In contrast, the spiked creatinine concentration ranged from 0 to 140 mg/dL, in increments of 7 mg/dL, and the spiked ethanol concentration ranged from 0% to 0.86%, in increments of 0.043%. The concentration errors introduced in making the samples were calculated to be less than 0.2%. Samples were stored in a refrigerator and used within 4 days after being made.

2.3 *Data Acquisition*

Spectra were acquired in both geometries, on successive days. Samples were warmed up to room temperature before spectral acquisition. The order of taking spectra was randomized and not correlated with either the target analyte or interferent's concentrations. The cuvette or the LCOF was first rinsed with 1–3 mL of sample before insertion of more of that same sample for measurement; this removed traces of the previous sample. For spectral acquisition, 3 mL of sample was used in the cuvette geometry and 1 mL was injected through the LCOF system. The actual volume of the LCOF tube itself is approximately 200 μL. Used sample was collected in a beaker and disposed of after the experiment. At the start of each measurement day, the system was aligned for maximum signal strength using a 20% ethanol aqueous solution. Spectra of all samples in a given dataset were taken within 1 day for one geometry, with no intentional change in laser power or alignment. The laser power fluctuation was less than 0.4%, as determined by monitoring the 880 cm^{-1} ethanol peak area using pure ethanol in the cuvette geometry integrated for 30 min.

Spectra were acquired with 1 s exposure time, and 200 such spectra were taken with no delay beyond the CCD readout time. Spectra corresponding to different effective integration times (from 1 to 200 s) could be created by averaging the proper number of acquisitions. There was a 5–10 s delay between filling the LCOF and starting data acquisition. After each injection of liquid into the LCOF, the system was observed by the operator through an infrared (IR) viewer. Locations of unusually high scattering loss within the LCOF, presumably due to microbubbles, appeared much brighter in the IR image. In such cases, the sample was injected again until the scattering was within a more typical range.

The spectra were preprocessed by background subtraction, cosmic ray removal, third-order Savitsky–Golay smoothing[21] (smoothing width$\approx$7 cm^{-1}), and approximate correction for the spectral throughput of the optical system (tungsten-halogen bulb spectrum). In urine spectra, the strong urea peak region from 973.5 to 1043.5 cm^{-1} was eliminated to remove its confounding effect upon concentration predictions. A few spectra whose intensities differed from the ensemble mean by more than 10% in the 800–900 cm^{-1} regime were eliminated as outliers; such spectra were readily detectable by eye as well. As a final step, a fifth-order polynomial was derived by least-squares fitting and subtracted from each spectrum to remove the broad spectral background.

Concentrations were predicted using a partial least squares (PLS)[22] leave-one-out cross validation implemented using in-house MATLAB code (The Mathworks, Natick, MA). In all cases, models of rank 3 or 4 were used. To investigate the relative importance of shot noise and other error factors, creatinine concentrations were predicted using sample sets corresponding to different effective integration times. For most calculations, the prediction set and the training set used the same integration time; an alternative approach is discussed later. The predicted concentrations were compared to the reference concentrations and the root mean squared error of cross validation (RMSECV) was calculated.

For two urine samples, we observed anomalously large prediction errors that were consistent between the two geometries. These samples were discarded *a posteriori* from further analysis, based upon the high likelihood of an error in the reference concentration. The final numbers of samples included in the analysis below were therefore 20 for DI/cuvette and 17 each for DI/LCOF, urine/cuvette, and urine/LCOF.

3 Results

3.1 *DI Water Samples*

PLS cross validation was used on spectral datasets from the two geometries to predict creatinine concentrations. The RMSECV was calculated for different effective integration times, as shown in the left-hand plot of Fig. 2. Error bars at most timepoints represent the standard deviation of RMSECVs obtained from different time windows; at the two rightmost timepoints, they are extrapolations.

Several observations can be made. At shorter integration times, the LCOF spectra produce lower prediction errors than the cuvette spectra. The errors from both geometries reduce as integration time increases, with the cuvette curve reducing more dramatically until it crosses the LCOF curve after 32 s. Both curves plateau at long integration times (note that the time axis is logarithmic), with a more pronounced flattening for the LCOF geometry.

Sample-by-sample creatinine predictions corresponding to two particular integration times are shown in Figs. 3 and 4. Figure 3 shows the predictions for 1 s spectra (short integration time) versus reference concentration, and Fig. 4 shows the corresponding predictions for 64 s spectra. For both ge-

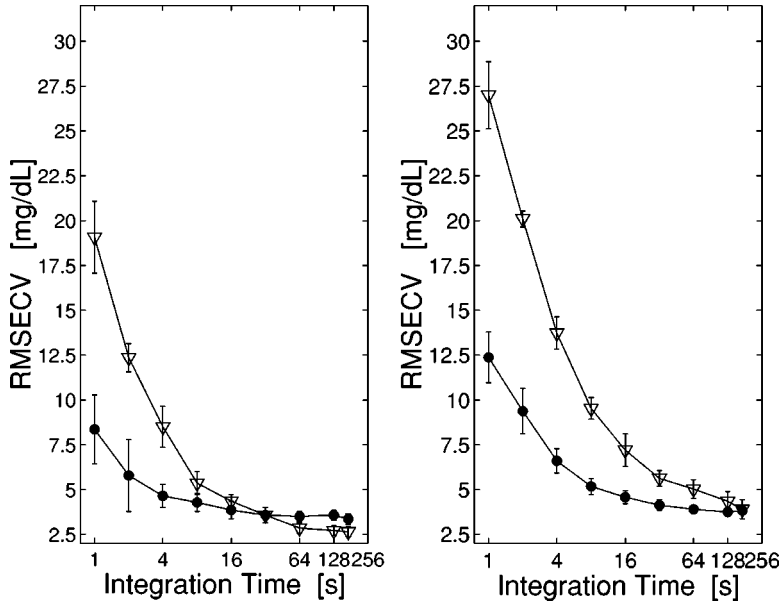

Fig. 2 Plot, versus effective integration time, of RMSECV for creatinine concentration predictions via partial least squares. Solid circles, LCOF data; open triangles, cuvette data. Left plot, results in deionized water; right plot, urine. Note the log scale of the *x* (time) axes.

ometries, 64 s spectra give better predictions than 1 s spectra. The cuvette prediction plot at 1 s exhibits a biased slope relative to the unity line.

The figures' insets show the major creatinine peak centered at 687.5 cm^{-1} as measured in the two geometries. At 1 s, when the peak is noisy, the LCOF has a higher signal to noise ratio (i.e., better spectral quality) than the cuvette, whereas at 64 s the peaks look equivalently smooth. The insets also illustrate the collection enhancement provided by the LCOF, which averaged 4.5 for the aqueous data set. The magnitude of this enhancement is discussed later.

3.2 *Urine Samples*

Typical Raman spectra of one urine sample, integrated for the first 64 s of spectral acquisition in each geometry, are shown in Fig. 5. The urea peak at 1014 cm^{-1}, the ethanol peak at 890 cm^{-1}, and the creatinine peak at 687.5 cm^{-1} are all stronger

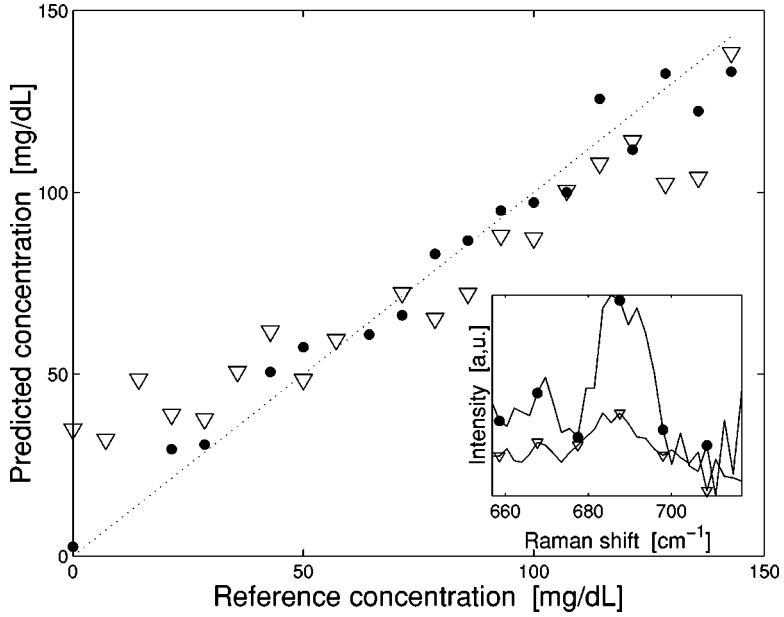

Fig. 3 Plot of Raman-based creatinine concentration predictions versus reference concentration in DI water when the effective integration time for calibration and prediction spectra is 1 s. Solid circles, LCOF data; triangles, cuvette data. The inset shows details of the creatinine Raman peak centered at 687.5 cm^{-1}. Circles, LCOF; triangles, cuvette.

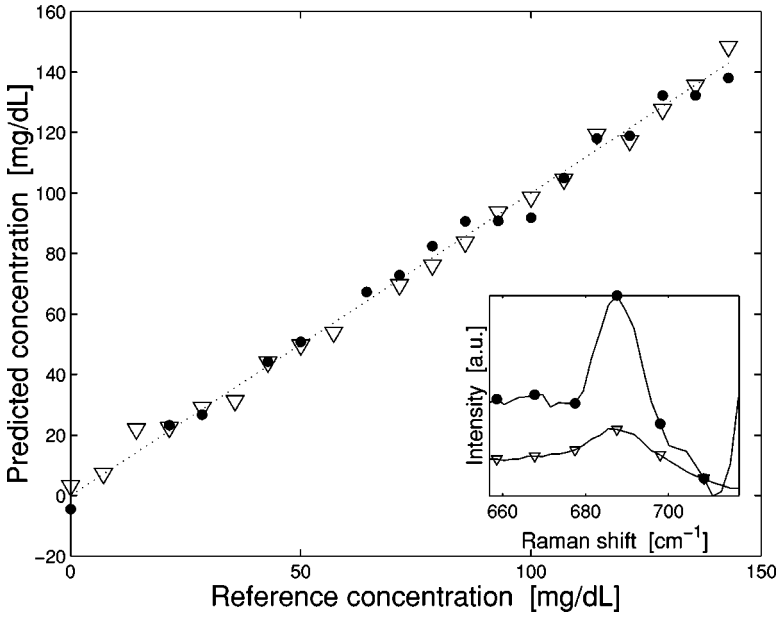

Fig. 4 Plot of Raman-based creatinine concentration predictions versus reference concentration in DI water when the effective integration time for calibration and prediction spectra is 64 s. Solid circles, LCOF data; triangles, cuvette data. The inset shows details of the creatinine Raman peak centered at 687.5 cm^{-1}. Circles, LCOF; trianges, cuvette. The LCOF peak area is greater by a factor of 4.5.

when measured in the LCOF geometry (upper) than in the cuvette (lower). Additional bands due to residual sampling of Teflon®-AF appear in the upper spectrum only.

The integrated spectra shown in Fig. 5 do not capture the fact that each specimen's broad spectral background decayed with continuing laser exposure, presumably due to fluorescence photobleaching. By subtracting a DI spectrum obtained using the same geometry, we could isolate this fluorescent component. Figure 6 shows the integrated fluorescence signal from 1100 to 1200 cm^{-1} versus time measured for a typical sample using both geometries, with each curve normalized to 1 at the start. In 200 s, the LCOF fluorescence drops to 33% of its original value and decays as a smooth exponential. The cuvette curve drops only to 80% of its starting value and fluctuates much more on the time scale of both seconds and minutes.

The righthand plot of Fig. 2 shows the RMSECV in urine as a function of effective integration time in the two geometries. As in the DI water case, the LCOF has a lower RMSEP value at shorter times and the cuvette makes a more rapid improvement with increased integration time. Here, however, the curves cross later, at the end of the measurement period

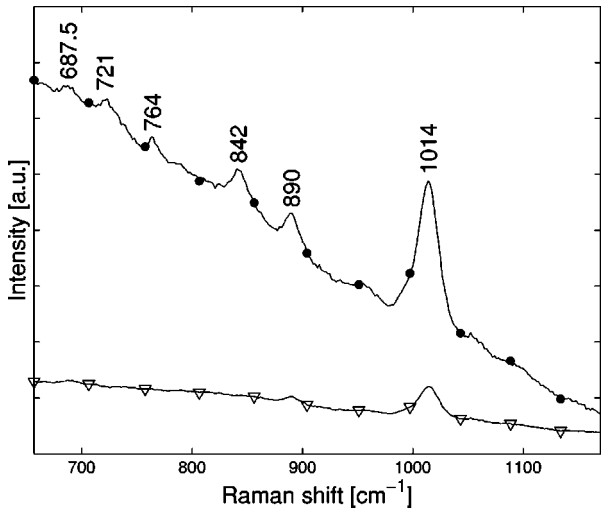

Fig. 5 Comparison of Raman peaks obtained using the two geometries. Upper curve (line with circles): LCOF-based spectrum of urine specimen spiked with ethanol and creatinine. Lower curve (line with triangles): cuvette-based spectrum of the same urine sample. Both spectra were integrated for 64 s. Spectra are base line-offset for clarity. The common peaks are seen to be much stronger in the LCOF spectrum. See text for discussion of the labeled peak assignments.

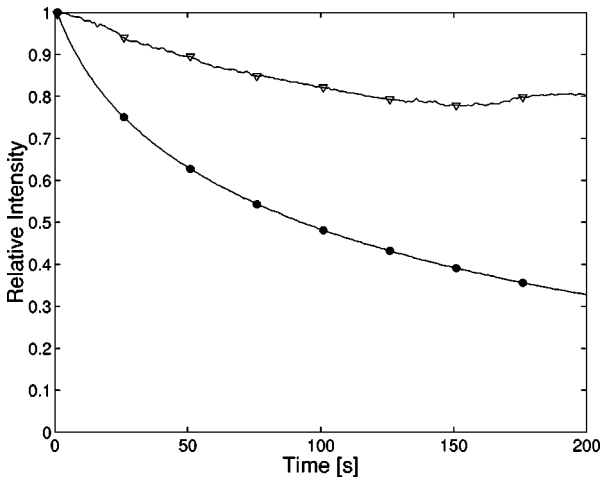

Fig. 6 Comparison of fluorescence photobleaching in urine specimens measured using the two geometries. Upper line with triangles, cuvette; lower line with circles, LCOF.

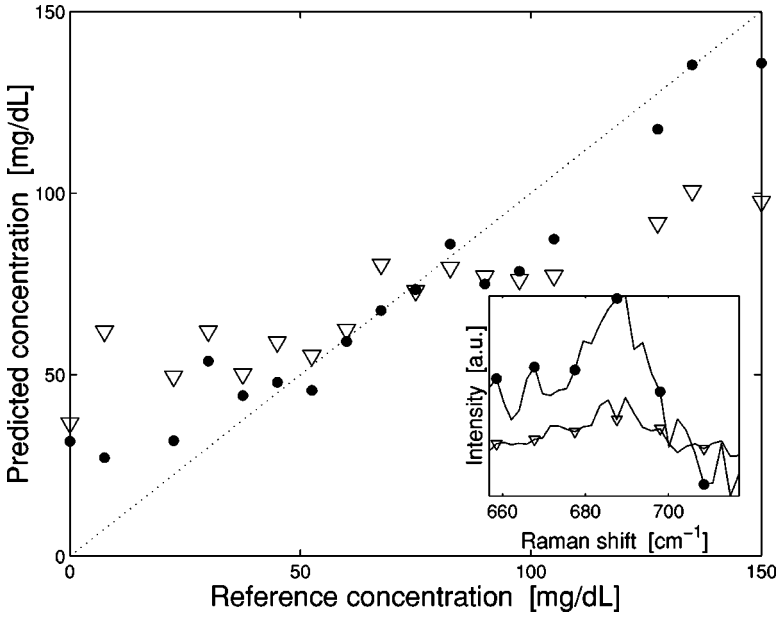

Fig. 7 Plot of Raman-based creatinine concentration predictions versus reference concentration in urine when the effective integration time for calibration and prediction spectra is 1 s. Solid circles, LCOF data; triangles, cuvette data. The inset shows details of the creatinine Raman peak centered at 687.5 cm^{-1}. Circles, LCOF; triangles, cuvette.

(176 s). Concentration predictions versus reference are shown for 1 s integration time and 64 s integration time in Figs. 7 and 8. Longer integration time improves the prediction for both geometries, and again the cuvette-based predictions at 1 s exhibit a biased slope relative to the unity line. The insets show the Raman peaks; in this case, the enhancement averaged 3.5.

4 Discussion

4.1 *Sources of Prediction Error*

Errors in concentration prediction can be grouped into three categories according to their origin: concentration error, random noise error and system error. The total RMSECV will be a quadratic sum of the prediction errors introduced by each noise source:

$$\mathrm{RMSECV}_{\mathrm{tot}}^2 = \mathrm{RMSECV}_{\mathrm{conc}}^2 + \mathrm{RMSECV}_{\mathrm{rand}}^2 + \mathrm{RMSECV}_{\mathrm{syst}}^2. \quad (1)$$

The three error categories are defined below.

Concentration error is generated in sample preparation. It characterizes how well the reference concentrations represent the true concentrations. In our experiments, this error was the same for both geometries because the same exact samples were used. We estimate our concentration error for creatinine at 0.2%.

Random noise error is caused by fluctuations in the recording of photon counts, leading to spectral noise and to subsequent error in concentration predictions. Sources of random error include Johnson noise, electrical line noise, and shot noise. In these experiments, the shot noise far exceeds the other random sources, due to the use of a cooled CCD detector with dedicated electronics. Since shot noise is random, its amplitude increases as the square root of the number of photon counts. If the LCOF geometry is M times more efficient than the cuvette geometry, the SNR for Raman peaks (defined here as signal divided by shot noise) will therefore be $\sqrt{M}$ times larger. We observe here, for later discussion, that for a fixed PLS calibration, $\mathrm{RMSECV}_{\mathrm{rand}}$ scales inversely with the SNR of the validation spectra. Given the SNR advantage of the LCOF system, this implies that

$$\frac{(\mathrm{RMSECV}_{\mathrm{rand}}^2)_{\mathrm{cuv}}}{(\mathrm{RMSECV}_{\mathrm{rand}}^2)_{\mathrm{LCOF}}} = M \quad (2)$$

in such cases.

System error is here defined as the error from various changes in the optical system, such as slight displacement of optical elements or the presence of microscopic bubbles in the samples, that vary from one sample measurement to the next. As discussed in Sec. 1, these instabilities affect the LCOF enhancement factor and also the spectral background. These variations cause prediction errors, and unlike random noise they are not mitigated by longer integration time.

Due to the greater instability of the present LCOF system, we expect LCOF predictions to have bigger system error than the cuvette predictions. At large integration times, as the random noise error decreases, we therefore expect the cuvette predictions to approach a lower error limit, generally, than those of the LCOF. We note that this is a first-generation LCOF system and that the LCOF system error can be reduced further in subsequent designs.

In the urine samples, there is an additional noise source: the variable fluorescence amplitude and decay versus time. As shown in Fig. 4, the photobleaching was both more extensive and more stable during the LCOF measurements, due to the illumination of the complete sample volume. This makes sense because in the LCOF geometry, the entire sample in the tubing (about 200 μL) is illuminated, as noted by others,[19] whereas in the cuvette geometry, only a small portion of the 3

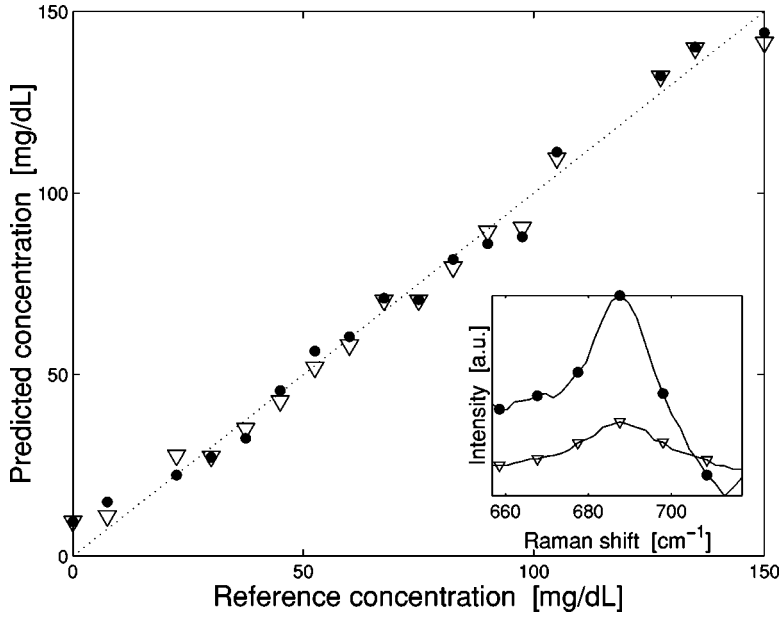

Fig. 8 Plot of Raman-based creatinine concentration predictions versus reference concentration in urine when the effective integration time for calibration and prediction spectra is 64 s. Solid circles, LCOF data; triangles, cuvette data. The inset shows details of the creatinine Raman peak centered at 687.5 cm^{-1}. Circles, LCOF; triangles, cuvette. The LCOF peak area is greater by a factor of 3.5.

mL sample is illuminated and photobleached, near the optical axis. In the latter case, passive exchange of molecules constantly brings new fluorophores into the region, maintaining the fluorescence at a relatively high level and introducing more temporal fluctuation. With regards to this noise source, then, the LCOF geometry should have an extra prediction advantage over the cuvette, seen only in the urine.

4.2 *Trends in Prediction Error*

The behavior of the RMSECV curves in Fig. 2 can be explained in terms of the error sources described earlier. In all cases, the RMSECV noticeably decreases versus effective integration time t; this indicates a significant level of random noise error, since that is the only type that decreases versus time. At the smallest values of t, random noise error is clearly dominant. Better predictions should therefore correspond to the spectra with better SNR. This is indeed the case: the LCOF predictions have the lower RMSECV at short times, and the LCOF spectra have the higher SNR due to the collection enhancement advantage of M.

In both plots, as integration time increases, the LCOF predictions reach a limiting value. This is because the random noise error has become small compared to the system error. Better predictions should now correspond to the spectra with lower system error. Again, this is indeed the case: the cuvette geometry has a lower RMSECV at longer integration times, and it has a lower level of system error in our current setup.

Notably, the crossing point in the RMSECV curves of Fig. 2 occurs much later in the urine than in the DI water. This is presumably due to the extra photobleaching advantage in urine, as mentioned earlier. Since most biological liquids will exhibit some fluorescence of this sort, it is important to note this additional advantage of the LCOF geometry.

For short integration times, in Figs. 3 and 7, there is an obvious bias in the slope of the predictions, particularly for the cuvette data. This is indicative of a weak calibration model, due to high levels of noise in the training data. The cuvette results show this effect more strongly because the SNR there is lower. For longer integration times, this bias is not observed because the SNR in both datasets has been improved by the $\sqrt{t}$ dependence.

4.3 *Quantitative Comparison*

The previous section described the qualitative trends in the RMSECV as collection time—and thus SNR—was systematically increased. A more quantitative analysis is also desirable. This is difficult to perform on the previous plots, however, because the RMSECV dependence upon SNR is complicated.

To make quantitative analysis simpler, the cross validations for the DI data were redone in an alternative fashion. As before, the effective integration time was systematically increased, but only for the validation spectra. Calibration spectra were now calculated using a fixed, large integration time (172 s), regardless of the SNR of the validation spectrum. In this case, where the PLS model is unchanging, $\mathrm{RMSECV}_{\mathrm{rand}}$ can be expressed as a function of time

$$\mathrm{RMSECV}_{\mathrm{rand}}(t) = \mathrm{RMSECV}_{\mathrm{rand}}(t_1)/\sqrt{t}, \tag{3}$$

where t is in seconds and $\mathrm{RMSECV}_{\mathrm{rand}}(t_1)$ is the amount of random error introduced into the predictions when the validation spectrum has an integration time of 1 s. Because the other noise sources do not depend upon integration time, Eq. (1) can be written as

$$\mathrm{RMSECV}^2_{\mathrm{tot}}(t) = \mathrm{RMSECV}^2_{\mathrm{rand}}(t_1)/t + \mathrm{RMSECV}^2_{\mathrm{other}}. \tag{4}$$

The results from the revised cross-validations for each geometry are shown as the solid points in Fig. 9. The corre-

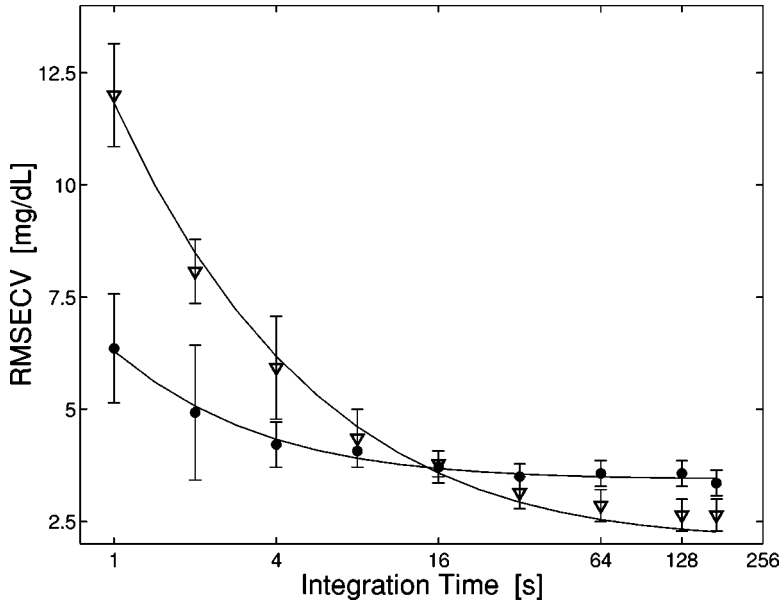

Fig. 9 Plot of RMSECV for DI water data, similar to Fig. ■ except that the effective integration time is varied for the prediction spectra only. As before: solid circles, LCOF data; triangles, cuvette data. The solid curves are the results of a least-squares fit to Eq. (■). Note the log scale of the *x* (time) axis.

sponding least-squares fits to Eq. (4) are plotted as lines. Values emerging from the fits are shown in Table 1. These values show that

$$\left[\frac{(\mathrm{RMSECV}_{\mathrm{rand}})_{\mathrm{cuv}}}{(\mathrm{RMSECV}_{\mathrm{rand}})_{\mathrm{LCOF}}}\right]^2 = 4.9 \approx M(=4.5), \qquad (5)$$

as predicted by Eq. (2). The SNR advantage provided by the LCOF geometry thus manifests itself in the expected way. The $\mathrm{RMSECV}_{\mathrm{other}}$ values also confirm that the present LCOF geometry introduces more system error than does the cuvette (3.4 versus 2.1 mg/dL).

We note that the value of $M=4.5$ for the creatinine peak, while reproducible for the many DI samples in this study, was small compared to other enhancements we observed even on the same measurement day. Samples of aqueous 20% ethanol showed enhancements of 8–10 for peaks at similar Raman shifts. We do not know if the relatively lower creatinine enhancement in water is due to the slightly higher absorption coefficient (water versus 20% aqueous ethanol), an unknown interaction with the LCOF material, or some effect that degrades the quality of the LCOF surface. We do notice that a particular piece of tubing irreversibly develops increased scattering loss over the scale of many days of experiments, leading to reductions in LCOF signal strength. In further studies we will coat the inner surface of the LCOF tubing to reduce this effect.[23]

Table 1 Values extracted from modeling the decrease of RMSECV versus integration time in Fig. 9. The RMSECV values are in units of mg/dL of creatinine.

Geometry	$\mathrm{RMSECV}_{\mathrm{rand}}$	$\mathrm{RMSECV}_{\mathrm{other}}$
LCOF	$5.3/\sqrt{t}$	3.4
Cuvette	$11.6/\sqrt{t}$	2.1

We also note that the enhancements using this LCOF at 830 nm excitation are small compared to literature reports using visible excitation. This is due to the significantly higher absorption coefficient of water in the 800–1000 nm regime.[24] We have previously measured ethanol peaks at different near-infrared emission wavelengths and shown that the LCOF enhancement quantitatively agrees with a theoretical model that includes the absorption spectrum of water.[20] Urine, unlike water, has significant absorption in the visible and an absorption minimum near 830 nm (data not shown), making the excitation wavelength more appropriate for such samples.

Analogous quantitative calculations were not possible with the urine spectra. Due to the time-dependent fluorescent background decay, spectra at different timepoints had different amounts of fluorescence signal. It was therefore not feasible to calibrate PLS using long-integration spectra and use that model on spectra from smaller time windows.

4.4 *Comparison of Results for Same Geometry, Different Medium*

Comparing the cuvette RMSECV results (Fig. 2) in different host liquids, we see that the curves have nearly the same shape. The errors in urine are, however, consistently higher than in the water samples. This is presumably because there are additional errors associated with urine than with water: more background chemicals (although nominally not changing from sample to sample in this study), additional fluorescence, and variations in fluorescence quenching.

The LCOF curves in the two host media also look similar. Again, the errors in urine are consistently higher than that in DI water. This time, however, the difference is less pronounced at longer integration times; i.e., the urine predictions are almost as accurate when system error is dominant. We speculate that this may be related to urine having higher absorption than water in this wavelength regime. Although the

higher absorption causes a lower enhancement (as was observed), it also shortens the mean sampling depth or ''effective length''[25] of the LCOF. Since there will be fewer irregularities (bubbles, rough spots) over a shorter length, the system error may be less when the absorption is higher, thereby offsetting the reduction in signal enhancement.

4.5 *Summary*

We have demonstrated concentration measurements of a biological analyte (creatinine) in two aqueous media (deionized water and urine) using Raman spectroscopy and a LCOF-based excitation-collection geometry. The enhanced signal collection, as compared to a standard cuvette geometry, was sufficiently stable to provide prediction accuracy of a few percent. For short integration times, the LCOF prediction error was lower than for the cuvette, due to the more favorable shot-noise-to-signal ratio.

For longer integration times, where shot noise is less of a factor, the LCOF prediction errors were higher than for the cuvette, due to the less stable geometry of this first-generation system. The crossover time (beyond which cuvette error becomes lower than LCOF error) was larger in urine than in DI; i.e., the LCOF had better relative performance in the more biological sample. This was due to two effects: more efficient and stable photobleaching of urine fluorescence, which favored the LCOF geometry over that of the cuvette; and a shorter effective LCOF collection length, which favored urine measurements over those in DI.

These studies suggest that accurate measurement of biological analytes at physiological concentrations in urine can be performed using an LCOF geometry. More accurate results are possible than those reported here; neither LCOF stability nor excitation wavelength were optimized in this pilot study. For eventual applications, the low sample volume requirement and efficient photobleaching effect in LCOFs make them an attractive measurement geometry for analyzing other clear biological liquids, including tear film, sweat, saliva, and blood serum.

Acknowledgments

The authors thank Heather Berkley for assistance in acquiring the Raman spectra. The kind permission of DuPont to use Teflon®-AF for this research project is gratefully acknowledged. This work is funded in part by a Biomedical Engineering Research Grant from the Whitaker Foundation.

References

1. M. J. Goetz, Jr., G. L. Coté, R. Erckens, W. March, and M. Motamedi, ''Application of a multivariate technique to Raman spectra for quantification of body chemicals,'' *IEEE Trans. Biomed. Eng.* **42**, 728–731 (1995).
2. J. P. Wicksted, R. J. Erckens, M. Motamedi, and W. F. March, ''Raman spectroscopy studies of metabolic concentrations in aqueous solutions and aqueous humor specimens,'' *Appl. Spectrosc.* **49**, 987–993 (1995).
3. X. Dou, Y. Yamaguchi, H. Yamamoto, S. Doi, and Y. Ozaki, ''Quantitative analysis of metabolites in urine using a highly precise, compact near-infrared Raman spectrometer,'' *Vib. Spectrosc.* **13**, 83–89 (1996).
4. F. H. M. Jongsma, R. J. Erckens, J. P. Wicksted, N. J. C. Bauer, F. Hendrikse, W. F. March, and M. Motamedi, ''Confocal Raman spectroscopy system for noncontact scanning of ocular tissues: an *in vitro* study,'' *Opt. Eng.* **36**, 3193–3199 (1997).
5. A. J. Berger, Y. Wang, and M. S. Feld, ''Rapid, non-invasive measurement of biological analytes by near-infrared Raman spectroscopy,'' *Appl. Opt.* **35**, 209–213 (1996).
6. A. J. Berger, T.-W. Koo, I. Itzkan, G. Horowitz, and M. S. Feld, ''Multicomponent blood analysis by near-infrared Raman spectroscopy,'' *Appl. Opt.* **38**, 2916–2926 (1999).
7. J. Y. Qu, B. C. Wilson, and D. Suria, ''Concentration measurements of multiple analytes in human sera by near-infrared laser Raman spectroscopy,'' *Appl. Opt.* **38**, 5491–5498 (1999).
8. H. P. Buschman, E. T. Marple, M. L. Wach, B. Bennett, T. C. B. Schut, H. A. Bruining, A. V. Bruschke, A. van der Laarse, and G. J. Puppels, ''*In vivo* determination of the molecular composition of artery wall by intravascular Raman spectroscopy,'' *Anal. Chem.* **72**, 3771–3775 (2000).
9. W. R. Premasiri, R. H. Clarke, and M. E. Womble, ''Urine analysis by laser Raman spectroscopy,'' *Lasers Surg. Med.* **28**, 330–334 (2001).
10. E. B. Hanlon, R. Manoharan, T.-W. Koo, K. E. Shafer, J. T. Motz, M. Fitzmaurice, J. R. Kramer, I. Itzkan, R. R. Dasari, and M. S. Feld, ''Prospects for *in vivo* Raman spectroscopy,'' *Phys. Med. Biol.* **45**, R1–R59 (2000).
11. G. W. Walrafen and J. Stone, ''Intensification of spontaneous Raman spectra by use of liquid core optical fibers,'' *Appl. Spectrosc.* **26**, 585 (1972).
12. P. Dress and H. Franke, ''A cylindrical liquid-core waveguide,'' *Appl. Phys. B: Lasers Opt.* **63**, 12–19 (1996).
13. L. Song, S. Liu, V. Zhelaskov, and M. A. El-Sayed, ''Application of liquid waveguide to Raman spectroscopy in aqueous solution,'' *Appl. Spectrosc.* **52**, 1364–1367 (1998).
14. M. J. Pelletier and R. Altkorn, ''Raman sensitivity enhancement for aqueous protein samples using a liquid-core optical-fiber cell,'' *Anal. Chem.* **73**, 1393–1397 (2001).
15. B. J. Marquardt, P. G. Vahey, R. E. Synovec, and L. W. Burgess, ''A Raman waveguide detector for liquid chromatography,'' *Anal. Chem.* **71**, 4808–4814 (1999).
16. R. Altkorn, I. Koev, and M. J. Pelletier, ''Raman performance characteristics of Teflon-AF 2400 liquid-core optical-fiber sample cells,'' *Appl. Spectrosc.* **53**, 1169–1176 (1999).
17. M. Holtz, P. K. Dasgupta, and G. Zhang, ''Small-volume Raman spectroscopy with a liquid core waveguide,'' *Anal. Chem.* **71**, 2934–2938 (1999).
18. R. Altkorn, I. Koev, R. P. V. Duyne, and M. Litorja, ''Low-loss liquid-core optical fiber for low-refractive-index liquids: Fabrication, characterization, and application in Raman spectroscopy,'' *Appl. Opt.* **36**, 8992–8998 (1997).
19. M. J. Pelletier and R. Altkorn, ''Efficient elimination of fluorescence background from Raman spectra collected in a liquid core optical fiber,'' *Appl. Spectrosc.* **54**, 1837–1841 (2000).
20. D. Qi and A. J. Berger, ''Quantitative analysis of Raman signal Enhancement from Aqueous Samples in Liquid Core Optical Fibers,'' *Appl. Spectrosc.* **58**, 1165–1171 (2004).
21. A. Savitzky and M. Golay, ''Smoothing and differentiating of data by simplifued least squares procedures,'' *Anal. Chem.* **36**, 1627–1639 (1964).
22. D. M. Haaland and E. V. Thomas, ''Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information,'' *Anal. Chem.* **60**, 1193–1202 (1988).
23. O. Inya-Agha, S. Stewart, T. Veriotti, M. L. Bruening, and M. D. Morris, ''Control of Teflon AF 2400 permeability in a liquid-core waveguide by an ultra-thin crosslinked polyamide coating,'' *Appl. Spectrosc.* **56**, 574–578 (2002).
24. L. H. Kou, D. Labrie, and P. Chylek, ''Refractive indexes of water and ice in the 0.65 μm to 2.5 μm spectral range,'' *Appl. Opt.* **32**, 3531–3540 (1993).
25. R. Altkorn, M. D. Malinsky, R. P. van Duyne, and I. Koev, ''Intensity considerations in liquid core optical fiber Raman spectroscopy,'' *Appl. Spectrosc.* **55**, 373–381 (2001).