Fall 2013

Comparison and combination of near-infrared and Raman spectra for PLS and NAS quantitation of glucose, urea and lactate

Yatian Sun
University of Iowa

COMPARISON AND COMBINATION OF NEAR-INFRARED AND RAMAN SPECTRA FOR PLS AND NAS QUANTITATION OF GLUCOSE, UREA AND LACTATE

by

Yatian Sun

A thesis submitted in partial fulfillment
of the requirements for the Master of
Science degree in Chemistry
in the Graduate College of
The University of Iowa

December 2013

Thesis Supervisor: Professor Mark A. Arnold

Graduate College
The University of Iowa
Iowa City, Iowa

CERTIFICATE OF APPROVAL

_____

MASTER'S THESIS

_____

This is to certify that the Master's thesis of

Yatian Sun

has been approved by the Examining Committee
for the thesis requirement for the Master of Science
degree in Chemistry at the December 2013 graduation.

Thesis Committee: _____
Mark A. Arnold, Thesis Supervisor

_____
Gary W. Small

_____
Amanda J. Haes

ACKNOWLEDGMENTS

I would like to thank everyone who has supported and inspired me during my graduate study.

At the very first, I would like to express my sincere thanks to my supervisor, Dr. Mark A. Arnold. Thank you for your immense support, encouragement, and inspiration throughout my graduate study. Your wisdom and knowledgeable discussions with me helps me a lot in my research. Your strict and abundance comments had a direct impact on the final version and quality of this thesis. Without you, it is impossible for me to accomplish my graduate study in the University of Iowa. I feel really grateful to have you to be my advisor. Special thanks also go to my committee members for their expert advice.

I would like to thank all my colleagues, who gave me so many valuable suggestions and encouragement. I am honored to express my sincere gratitude to Dr. Jooyoung Choi, your help with the MATLAB codes and useful suggestions you gave me really inspired me in my research; and Dr. Min Ren, your data collection and well organized data files. I would also like to thank the rest members in Arnold's group for your support and help.

At last but not least, I would like to give my deepest gratitude to my family for their great support during my graduate studying.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Diabetes

Diabetes is a metabolic disorder characterized by abnormal blood glucose levels. For people with type I diabetes, the pancreas cannot provide enough insulin to convert glucose into energy.[1] For people with type II diabetes, insulin is produced but the body cannot use it properly. The World Health Organization (WHO) estimates 347 million people worldwide had some form of diabetes in 2012.[2]

According to the American Diabetes Association (ADA), poor control of blood glucose concentrations contributes to heart disease, high blood pressure, blindness, kidney failure and other chronic medical disorders.[3] Studies have shown that tight glycemic control can reduce the onset of these complications, thus improving the quality of life and extending the life expectancy for people with type I diabetes for up to eight years.[4] Tight glycemic control demands frequent blood glucose measurements coupled with a sophisticated insulin delivery algorithm.

Conventionally, individual glucose measurements are achieved by drawing a drop of blood and applying it to a glucose test strip within a glucose meter. This test-strip approach is invasive, painful, expensive and inconvenient.[5] Spectroscopic methods promise painless measurements that can be continuous, thereby providing an attractive means to track glucose, decrease complications, and improve the quality of life for people with diabetes.[6]

Several approaches are proposed for noninvasive glucose sensing, such as near-infrared (NIR) spectroscopy,[7] Raman scattering spectroscopy,[8,9] glucose diffusion analysis[10] and photo-acoustic spectroscopy.[11] The idea of these noninvasive methods is to transmit light through human skin, and collect a spectrum.[12,13] The spectrum is analyzed to provide the concentration of glucose. In NIR spectroscopy, glucose concentrations are

monitored by transmittance spectroscopy over the combination region of the near infrared spectrum (2.0 to 2.5 μm).[14] However, there are several factors that affect measurement accuracy, including different levels of fat and proteins within the skin matrix, environmental changes and long-term stability of the instrumentation. Raman spectroscopy measures inelastic light scattered molecules within the tissue matrix. Studies have shown that glucose concentrations are well correlated with Raman scatter intensities within various samples, such as blood, serum and water.[15,16,17,18,19] Noninvasive Raman measurements are complicated by fluorescence background signals in tissue samples and instability of the excitation laser.[19]

Near-Infrared Spectroscopy

The NIR region of the electromagnetic spectrum is associated with combination and overtone vibrations of C-H, O-H and N-H bonds.[20] The full range of the NIR spectrum spans over 12800-4000 cm^{-1} wavenumbers or 780-2500 nm wavelengths. Absorption of NIR radiation follows the Beer-Lambert law, presented in Equation I-1,

$$A = -\log\left(\frac{I}{I_0}\right) = \varepsilon \cdot b \cdot c \qquad \text{Equation I-1}$$

where A is the absorbance of a sample, I and I_0 are the intensities of the transmitted and incident radiation, respectively, ε denotes the absorptivity at a certain wavelength, b represents the optical path length, and c is the concentration of the sample.

Absorptivities for NIR absorption bonds can be one to three orders of magnitudes weaker than those in the mid-infrared region. Although lower absorptivities correspond to lower sensitivities, they also permit measurements in thicker samples, thereby permitting noninvasive measurements without sample preparation. In addition, NIR radiation is low energy and nondestructive to the sample. A principal drawback of NIR spectroscopy is the fact that NIR absorption bands are broad and overlapping, which demands

multivariate methods of analysis for data processing. For the work described here, the spectral region of interest is 5000 to 4000 cm^{-1}. This spectral range corresponds to an absorption minimum between two large water absorption bands centered at 5200 and 3600 cm^{-1}.[21]

A Fourier transform (FT) instrument is used to collect the NIR spectra for the work described in this thesis. The centerpiece of an FT spectrometer is the Michelson interferometer,[22] which is shown schematically in Figure I-1. A beam splitter separates the incident radiation into two paths with nearly equal power, one path is reflected and the other is transmitted. The two beams then recombine at the beam splitter after reflection by fixed and moving mirrors. Depending on the position of the moving mirror, the optical path length difference of the two beams creates an interference pattern recorded as an interferogram. The fast Fourier transform (FFT) algorithm is applied to convert the time-domain interferogram into a frequency-domain single-beam spectrum.[23]

An FT spectrometer has several advantages over dispersive systems.[24] In the FT configuration, radiation at all frequencies reaches the detector simultaneously, thus decreasing the acquisition time. More scans can be collected in the same time period which enhances the measurement signal-to-noise ratio (SNR). In addition, the absence of a slit enables high radiant powers at the detector, which also increases the SNR. A HeNe laser is used as the internal frequency standard to provide excellent frequency precision ($<\pm 0.01$ cm^{-1}) and excellent signal averaging.

Raman Scattering Spectroscopy

Raman scattering is a type of inelastic scattering in which the wavelength of the scattered photons is different than the wavelength of the incident light. The difference corresponds to vibrational modes within molecules responsible for the scattering. The basic mechanism of Raman scattering is shown in Figure I-2.

At room temperature, most molecules are in the ground electronic state.

Absorption of a photon whose energy does not fit any energy level within the molecule can excite the molecule to a virtual state. The energy of the virtual state is determined by the energy of the excitation photon. The excited molecule then relaxes back to the ground state. If the emitted radiation has the same energy as the excitation radiation, this is called Rayleigh scattering. The excited molecule can also relax back to a vibrational state within the ground electronic state. If the energy of the emitted photon is less than the excitation photon, this is called Stokes scattering. If the energy of the emitted photon is higher than the excitation photon, this is called anti-Stokes scattering. The probability for the occurrence of the anti-Stokes scattering is much lower than that of the Stokes scattering because the fraction of molecules in a higher vibrational energy level is small at room temperature.

For the case where the incident radiation has a frequency of $\nu_{ex,}$the electric field vector E of this radiation is expressed by Equation I-2 below:[25]

$$E = E_0 \cos(2\pi\nu_{ex} t) \qquad \text{Equation I-2}$$

where E_0 is the amplitude of the incident radiation wave. When the electric field interacts with an electron cloud of a molecular vibrational bond, it induces a dipole moment m in the bond and the dipole moment can be described by Equation I-3:

$$m = \alpha E = \alpha E_0 \cos(2\pi\nu_{ex} t) \qquad \text{Equation I-3}$$

where α is the polarizability of the bond and describes the deformability of the bond in an electric field. Equation I-4 specifies the dipole moment of Raman active mode within a molecule:

$$\mathrm{m} = \alpha_0 E_0 \cos(2\pi\nu_{ex} t) + \frac{E_0}{2} r_m \left(\frac{\partial\alpha}{\partial r}\right) \cos[2\pi(\nu_{ex} - \nu_v)t] + \frac{E_0}{2} r_m \left(\frac{\partial\alpha}{\partial r}\right) \cos[2\pi(\nu_{ex} + \nu_v)t] \qquad \text{Equation I-4}$$

where α_0 is the polarizability of the bond at the equilibrium internuclear distance, r_m is the maximum internuclear separation relative to the equilibrium position, and *r* is the internuclear separation at any instant. The first term in Equation I-4 represents Rayleigh scattering at the excitation frequency ν_{ex}, the second term corresponds to the Stokes scattering at the frequency $\nu_{ex} - \nu_v$, and the third term represents the anti-Stokes scattering at the frequency $\nu_{ex} + \nu_v$.

A change of polarizability is required for Raman scattering to take place. In contrast, a change of dipole moment is required for absorption of infrared frequencies. In infrared absorption, only radiation of the same frequency as the molecular vibration interacts with the molecule and promotes it to an excited vibrational state. In Raman scattering, the molecule is promoted and reemits the radiation and returns to a different vibrational level.

Raman scattering spectroscopy is typically used for structure determination. Quantitative analysis with Raman spectroscopy has also been described in the literature.[26,27] Equation I-5 below relates the intensity of Raman scattered light $I(\nu)_R$ to various parameters:

$$I(\nu)_R = \frac{2^4\pi^3}{45\cdot 3^2\cdot c^4}\cdot\frac{hI_L N(\nu_0-\nu)^4}{\mu\nu\left(1-e^{-h\nu/kT}\right)}\cdot[45(\alpha'_a)^2+7(\gamma'_a)^2] \qquad \text{Equation I-5}$$

where *c* is the speed of light, *h* is the Plank's constant, I_L denotes the excitation intensity, *N* is the number of scattering molecules, ν represents the molecular vibrational frequency in Hertz, ν_0 is the laser excitation frequency in Hertz, μ is the reduced mass of the vibrating atoms, *k* is the Boltzmann constant, *T* is the absolute temperature, α'_a is the mean value invariant of the polarizability tensor, and γ'_a is the anisotropy invariant of the polarizability tensor. Equation I-5 shows that the intensity of Raman scattered light is proportional to the number of molecules *N*, which defines the concentration dependency assuming a constant volume of sample is probed in the measurement. The

correlation between Raman intensity and concentration of analyte is the basis for quantitative analysis using Raman spectroscopy.[28,29]

Both Raman spectroscopy and NIR spectroscopy are related to the vibrational modes of the molecules, however, due to different mechanisms, some vibrational modes in one molecule are Raman active while some vibrational modes are IR active, and some vibrational modes are both Raman and IR active. This provides complementary features between Raman and NIR measurements.

Partial Least-Square (PLS) Regression

Absorption features at NIR frequencies are broad, weak and highly overlapping, thereby requiring multivariate analysis methods for quantitative analysis.[30,31] Multivariate calibration techniques help to extract information for an analyte of interest from within a complex spectral matrix. In calibration, the relationship between spectral variances and changes in the concentration of the target analyte is established. The multivariate algorithms used in this dissertation are partial least squares (PLS) regression and net analyte signal (NAS) analysis.

The PLS regression method decomposes the spectral and the concentration variances into loadings and scores and correlates the spectral variance with the concentration variance.[32] PLS is a powerful and reliable tool to analyze spectral data because it extracts latent variables by taking into account covariance of both the spectral and concentration matrices.[33,34]

The PLS algorithm works as follows:

The first loading weight $\boldsymbol{w}_1$ is calculated according to Equation I-6,

$$\boldsymbol{w}_1 = \frac{\boldsymbol{R}^T\boldsymbol{c}}{\|\boldsymbol{R}^T\boldsymbol{c}\|} \qquad \text{Equation I-6}$$

where $\boldsymbol{R}$ is the n$\times$p spectral matrix, and c is the n$\times$1 vector of analyte concentration.

The first score $\boldsymbol{t}_1$ is calculated by Equation I-7,

$$\boldsymbol{t}_1 = \boldsymbol{R}\boldsymbol{w}_1 \qquad \text{Equation I-7}$$

The first spectral loading is calculated by Equation I-8,

$$\boldsymbol{p}_1 = (\boldsymbol{t}_1^T\boldsymbol{t}_1)^{-1}\boldsymbol{R}^T\boldsymbol{t}_1 \qquad \text{Equation I-8}$$

and the corresponding first concentration loading is calculated by Equation I-9,

$$\boldsymbol{q}_1 = (\boldsymbol{t}_1^T\boldsymbol{t}_1)^{-1}\boldsymbol{t}_1^T\boldsymbol{c} \qquad \text{Equation I-9}$$

Before the second latent variable is calculated, both the spectral matrix and concentration vectors must be recalculated by Equations I-10 and I-11:

$$\boldsymbol{R} = \boldsymbol{R} - \boldsymbol{t}_1\boldsymbol{p}_1^T \qquad \text{Equation I-10}$$

$$\boldsymbol{c} = \boldsymbol{c} - \boldsymbol{t}_1\boldsymbol{q}_1 \qquad \text{Equation I-11}$$

The algorithm is repeated with the new $\boldsymbol{R}$ and $\boldsymbol{c}$ values to get another set of $\boldsymbol{w}$, $\boldsymbol{t}$, $\boldsymbol{p}$, $\boldsymbol{q}$ and the process is repeated for the number of specified factors. After all the desired factors, or latent variables, are calculated, the regression coefficient b is calculated as follows:

$$\boldsymbol{b} = \boldsymbol{W}(\boldsymbol{P}^T\boldsymbol{W})^{-1}\boldsymbol{q} \qquad \text{Equation I-12}$$

where $\boldsymbol{W}$ is a p$\times$k matrix with k weighting factors ($\boldsymbol{w}_1$ to $\boldsymbol{w}_k$), $\boldsymbol{P}$ is a p$\times$k matrix with k spectral loadings ($\boldsymbol{p}_1$ to $\boldsymbol{p}_k$), and $\boldsymbol{q}$ is the concentration loading vector. Concentration predictions can then be calculated by:

$$\hat{\boldsymbol{c}} = \boldsymbol{R}\boldsymbol{b} \qquad \text{Equation I-13}$$

To evaluate the performance of the calibration model, the standard error of calibration (SEC) and the standard error of prediction (SEP) are calculated as shown

below: [35,36]

$$SEC = \sqrt{\frac{\sum(\boldsymbol{c}_c - \hat{\boldsymbol{c}}_c)^2}{n_c - k - 1}} \qquad \text{Equation I-14}$$

$$SEP = \sqrt{\frac{\sum(\boldsymbol{c}_p - \hat{\boldsymbol{c}}_p)^2}{n_p}} \qquad \text{Equation I-15}$$

where $\boldsymbol{c}_c$ and $\boldsymbol{c}_p$ are the known concentrations of the analyte in the calibration and prediction data sets, respectively, and $\hat{\boldsymbol{c}}_c$ and $\hat{\boldsymbol{c}}_p$ are the concentration predictions from the calibration model for the calibration and prediction data sets, respectively. n_c and n_p are the number of spectra in the calibration and prediction data sets, respectively, and k is the number of latent variables used in the PLS model.

Net Analyte Signal (NAS)

Net analyte signal (NAS) is defined as the part of a sample spectrum that is orthogonal to the spectra of all other components in the sample, and is, therefore, the unique spectral signature for the analyte of interest. The NAS method was originally proposed by Lorber in 1986.[37,38] Figure I-2 is a representation of the NAS vector that is orthogonal to the background. The NAS vector can be computed by the following equation:

$$\boldsymbol{NAS}_i = (\boldsymbol{I} - \boldsymbol{A}_{-i}\boldsymbol{A}_{-i}^{+})\boldsymbol{a}_i \qquad \text{Equation I-16}$$

where $\boldsymbol{I}$ is the identity matrix, $\boldsymbol{A}_{-i}$ is the background matrix containing all sources of variance except the target analyte, $\boldsymbol{A}_{-i}^{+}$ represents the pseudo inverse of $\boldsymbol{A}_{-i}$, $\boldsymbol{a}_i$ is the pure component spectrum of the target analyte.

The background matrix should include spectral variance associated with the blank and spectral features of the chemical components within the sample matrix. The spectral variance of the background spectra can be characterized by a principal component

analysis (PCA).[38] In this method, a selected number of principal components (PCs) is used to represent the background spectral variance while removing noise. The NAS is computed by projecting the pure component spectrum onto the first PC. The remaining orthogonal spectrum is then projected onto the second PC and this process continues for all PCs. The equation of the projection onto the k^{th} PC is shown below:

$$\boldsymbol{a}_i^* = \boldsymbol{a}_i - \frac{\boldsymbol{a}_i \cdot \boldsymbol{v}_k}{\boldsymbol{v}_k^T \cdot \boldsymbol{v}_k} \boldsymbol{v}_k \qquad \text{Equation I-17}$$

where $\boldsymbol{a}_i^*$ is the remaining orthogonal spectrum and $\boldsymbol{v}_k$ is the k^{th} PC. The final $\boldsymbol{a}_i^*$ after projecting onto each PC is termed the NAS.

The NAS calibration coefficient can be calculated by the following equation:

$$\boldsymbol{\beta} = \frac{\boldsymbol{NAS}}{\|\boldsymbol{NAS}\|^2} \qquad \text{Equation I-18}$$

where $\boldsymbol{\beta}$ is the calibration vector of the analyte of interest, $\boldsymbol{NAS}$ represents the NAS vector in the sample matrix and $\|\boldsymbol{NAS}\|$ denotes the magnitude of the NAS vector.

Selectivity is defined as the ability to quantify the analyte of interest accurately within the sample matrix. Selectivity in a NAS calibration model can be expressed as the ratio between the norm of the $\boldsymbol{NAS}_i$ vector and the norm of the pure analyte spectrum $\boldsymbol{a}_i$, as shown in the equation below:[37,38]

$$SEL = \frac{\|\boldsymbol{NAS}_i\|}{\|\boldsymbol{a}_i\|} \qquad \text{Equation I-19}$$

Sensitivity is defined as the slope of the calibration curve, which describes the change of the magnitude in signal with the change of concentration. In an NAS calibration model, the sensitivity can be calculated by Equation I-20:[37,38]

$$SEN = \frac{\|\boldsymbol{NAS}_i\|}{\boldsymbol{c}_i} \qquad \text{Equation I-20}$$

The SNR of the NAS model is defined as the ratio of the magnitude of the NAS vector to the noise. The noise is calculated as the standard deviation of the magnitude of the background spectral projection onto the NAS vector. The SNR in an NAS model can be expressed as:

$$\left(\frac{S}{N}\right)_{NAS} = \frac{\|NAS\|}{\sigma_{NAS}} \qquad \text{Equation I-21}$$

where σ_{NAS} is the noise along the NAS vector.

Overview of Thesis

Both NIR absorption spectroscopy and Raman scattering spectroscopy have the potential to provide noninvasive measurements of blood glucose levels *in vivo* to get frequent glucose readings and thereby enhance glycemic control. Real spectral data were analyzed and compared for NIR and Raman separately. The results show good prediction abilities for NIR and Raman individually.[21,39] The objective of this research is to compare the prediction performance of the PLS and NAS models based on NIR spectra, Raman spectra and the combination of NIR and Raman spectra for glucose, urea and lactate. The primary research question can be formulated as follows: Can analytical performance be improved by using a combination of NIR and Raman spectra?

In Chapter II, real NIR and Raman spectral data were used to evaluate whether combining the two techniques improves the prediction ability of the PLS model. The described research indicates that the SNR of these two spectral data sets are different, and adding low SNR spectra lessens the analytical performance of the PLS model. In Chapter III, simulated NIR and Raman spectral data are used to illustrate how concatenation of NIR and Raman spectra with the same NAS length and SNR improves prediction ability of PLS models for glucose, urea and lactate.

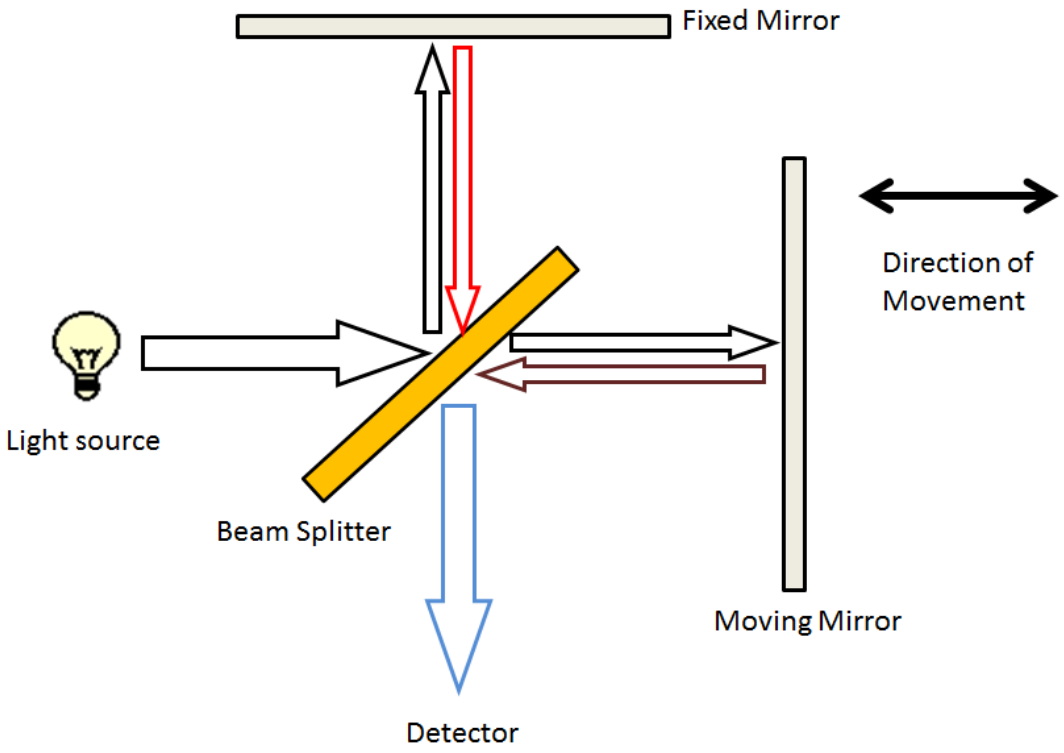

Figure I-1. Schematic diagram of a Michelson interferometer.

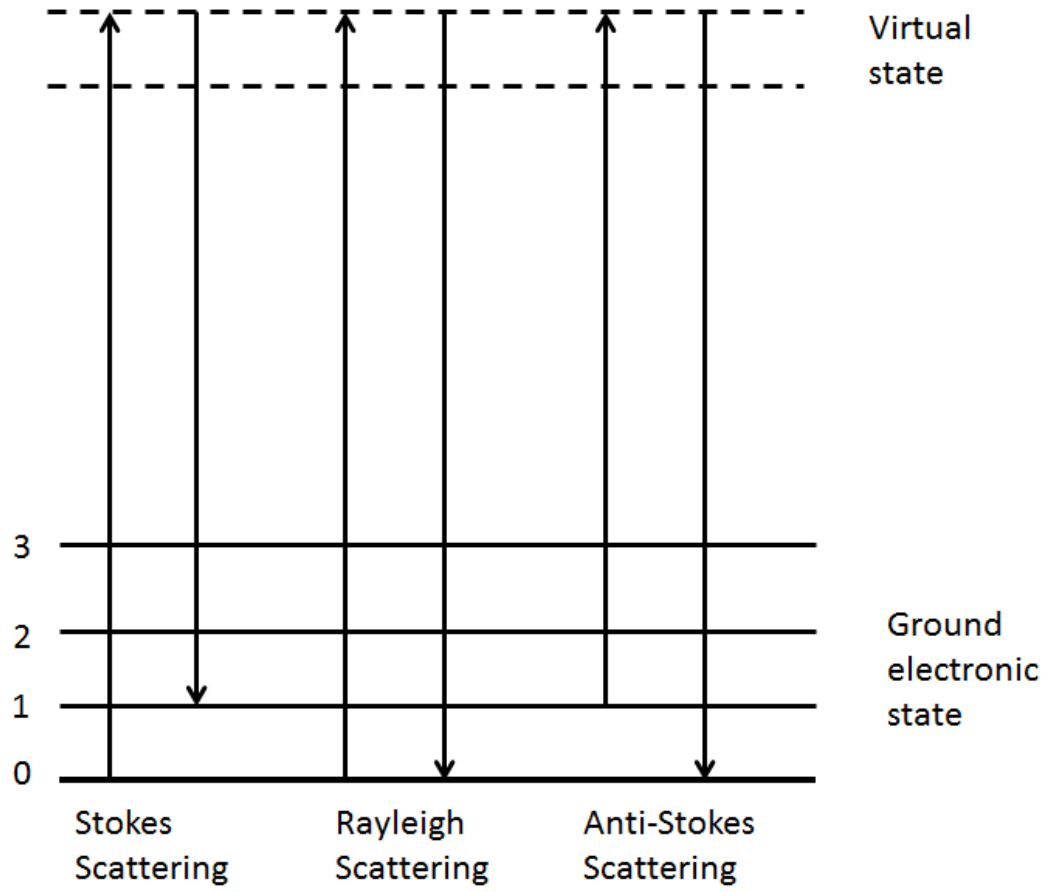

Figure I-2. Schematic diagram of Rayleigh scattering and Raman scattering.

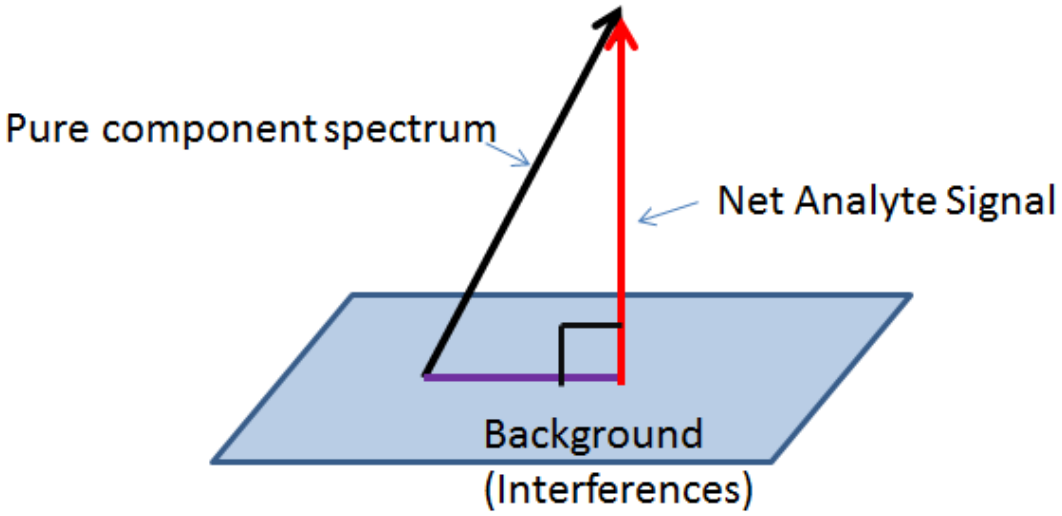

Figure I-3. Schematic diagram of NAS model.

CHAPTER II

COMPARISON AND COMBINATION OF NEAR-INFRARED AND RAMAN SPECTRA FOR PLS AND NAS QUANTITATION OF GLUCOSE, UREA AND LACTATE USING REAL DATA

Introduction

As discussed in Chapter I, tight glycemic control for people with type I or type II diabetes is known to delay the onset of complication caused by chronic hyper-glycemia.[4] Noninvasive spectroscopic measurements of glucose in people with diabetes is one strategy being pursued to enable the frequent monitoring necessary to achieve tight control.[40,11] Near infrared absorption spectroscopy and Raman scattering spectroscopy are the most promising methods reported to date for such noninvasive measurements.[8]

The objective of the research described in this chapter is to explore the analytical benefits of combining information derived from both NIR and Raman spectra for the purpose of measuring glucose. To the extent the information embedded within the NIR and Raman spectra is complementary, the hypothesis is that combining NIR and Raman spectra will provide improved multivariate calibration models for measuring the concentration of glucose.[41]

Although NIR spectroscopy and Raman spectroscopy are both associated with molecular vibrations, significant differences are noteworthy in terms of their complementary nature. First, NIR absorptions are associated with combinations and overtones of fundamental vibration modes associated with C-H, N-H and O-H bonds.[20] Raman spectroscopy, on the other hand, provides spectral features associated with fundamental vibrational modes for a wider range of molecular bonding.[42] In the work presented here, the NIR information is collected over the 4000-5000 cm^{-1} spectral range corresponding to combinations of stretching and bending vibrations. The Raman spectra are collected over the 700-1700 cm^{-1} spectral range, corresponding to fundamental modes

associated with C-H, C-O, C-C, and C-N molecular bonds.[42]

In addition, NIR spectroscopy measures absorption of electromagnetic (EM) radiation while Raman spectroscopy is based on scattering of EM radiation. The different selection rules for these two phenomena, changing dipole moment for NIR and changing polarizability for Raman, provide additional complementary information. Other differences include sensitivity to water and sensitivity to surfaces and interfaces. The complementary nature of these differences might serve as the basis for improving analytical performance.

The analytical utility of combining NIR and Raman spectra is evaluated here by concatenating NIR and Raman spectral data collected from a set of 60 samples. Multivariate PLS and NAS calibration models are generated from these concatenated NIR-Raman spectra and the performance of these models are compared to analogous models based on NIR and Raman spectra alone.

Experimental Section

The spectral data used in this study were collected previously and details are provided elsewhere.[21,39] Briefly, 60 samples were prepared with randomized concentrations of glucose, urea and lactate over the 1-30 mM concentration range. Solutions were prepared in a pH 6.86 phosphate buffer. Nine pure component NIR and Raman spectra were available for each component with concentrations ranging from 3 to 27 mM with a 3 mM increment.

NIR spectra were collected with a Nicolet 670 Nexus Fourier transform spectrometer (Nicolet Analytical Instruments, Madison, WI) equipped with a 20 watt tungsten-halogen lamp, a calcium fluoride beam splitter and a cryogenically cooled InSb detector. A K-band interference filter was set before the sample to restrict the spectral range to 4000-5000 cm^{-1}. Samples were placed in a water-jacketed thermostatted cell to maintain the temperature at 37 ± 0.1 ℃. Raman spectra were collected with a Mark II

holographic fiber-coupled probe head attached to a Hololab 5000R modular research Raman spectrograph (Model HoloSpec f/1.8i, Kaiser Optical System, Inc.) with a 785 nm excitation laser. For the Raman spectra, samples were placed in a 10 mm thick quartz cuvette and the collected spectral region was 100-3450 cm^{-1}. No temperature control was used for the sample and the lab temperature averaged 22.0 $\pm$ 0.3 ℃ during data collection.

For both the NIR and Raman data sets, raw spectra were collected in triplicate for each sample and buffer spectra were collected after every fifth sample. The most recently collected NIR buffer spectra were used as the reference for absorbance calculations and the mean of the Raman buffer spectra was used for background subtraction for each sample Raman spectrum. All data processing was performed using MATLAB (Math Works, Inc., Natick, MA).

Results and Discussion

The NIR absorbance spectra were calculated by taking the negative base-10 logarithm of the ratio of a sample single-beam spectrum and the single-beam spectrum of the corresponding phosphate buffer solution. Example absorption spectra are provided in Figure II-1 for the 27 mM pure component solutions of glucose, urea, and lactate. These NIR spectra reveal absorption bands centered at 4300, 4400 and 4700 cm^{-1} for glucose which correspond to the combination vibration of C-H and O-H bonds.[35] The absorption bands centered at 4575 and 4650 cm^{-1} for urea correspond to the combination vibrations of N-H bonds.[35] The absorption bands centered at 4350, 4450 and 4800 cm^{-1} for lactate correspond to combination vibrations for C-H and O-H bonds.[35] These absorption bands overlap to some extent. Negative absorbance values are created by a combination of slight variations in the spectrometer between sample and buffer spectra as well as water displacement effects over frequencies where the absorptivity of water is greater than that of the solutes.[43]

The Raman spectra were calculated by subtracting the corresponding averaged buffer spectra from each sample spectrum and examples are presented in the lower panel of Figure II-1 for the 27 mM pure component solutions of glucose, urea, and lactate. In comparison to the NIR spectra, features in the Raman spectra are narrower with less overlap. Glucose bands centered at 880, 1072, 1128 and 1370 cm^{-1} correspond to C-C stretching, C-O stretching, C-O-H bending and CH_2 wagging, respectively.[44,45] Urea bands centered at 1003 and 1160 cm^{-1} correspond to N-C-N stretching and NH_2 rocking, respectively.[46,47] Lactate bands centered at 855, 930, 1040, 1420 and 1455 cm^{-1} correspond to C-COO^- stretching, CH_3 rocking , C-CH_3 stretching, COO^- stretching and CH_3 deformation, respectively.[48,49]

Spectral Quality

Analysis of spectral ratio noise was used to compare spectral quality across the two fundamentally different types of spectra: NIR absorption spectroscopy measured from 4000-5000 cm^{-1} (2000-2500 nm) and Raman scattering spectroscopy measured with an excitation wavelength of 785 nm. In both cases, each raw single-beam spectrum collected for a sample was ratioed to the other two spectra collected for that sample. The root-mean-square (RMS) noise was calculated across each 100 cm^{-1} segment of the resulting ratioed spectra. The RMS noise is computed relative to a second order polynomial fit of the ratioed spectral data for each 100 cm^{-1} segment.

Figure II-2 presents a set of representative ratioed spectra across the full spectral ranges for both the NIR and Raman data sets and the measured RMS noise levels for each 100 cm^{-1} segment are summarized in Table II-1. For the NIR spectra, the ratioed spectra show how the RMS noise varies across the spectral range in accordance with the transmission properties of water. The lowest RMS noise levels correspond to frequencies where the water absorption is lowest (4400-4600 cm^{-1}), thereby providing highest radiant powers at the detector and, correspondingly, the highest SNR. For the Raman spectra, the

Table II-1. RMS noise level of ratioed spectra for NIR and Raman spectra.

Raman Spectral Range (cm^{-1})	Raman Noise Level (AU)	NIR spectral Range (cm^{-1})	NIR Noise Level (AU)
700-800	1.72×10^{-3}	4200-4300	9.46×10^{-5}
800-900	1.84×10^{-3}	4300-4400	1.57×10^{-5}
900-1000	2.02×10^{-3}	4400-4500	7.21×10^{-6}
1000-1100	2.30×10^{-3}	4500-4600	7.65×10^{-6}
1100-1200	2.67×10^{-3}	4600-4700	1.10×10^{-5}
1200-1300	2.85×10^{-3}	4700-4800	3.04×10^{-5}
1300-1400	2.95×10^{-3}		
1400-1500	3.33×10^{-3}		
1500-1600	3.58×10^{-3}		
1600-1700	3.45×10^{-3}		

ratioed spectra reveal less structure and are characterized by increasing RMS noise levels from low to high frequencies. Compared to the NIR ratioed spectra, RMS noise levels for the Raman spectra are greater in magnitude but more constant across the spectral range.

Figure II-3 shows the distribution of RMS noise levels determined for each sample in the two data sets. In both cases, the noise levels are uniformly distributed across the samples with no significant shifts or systematic variations in the RMS noise. The consistency in RMS noise levels across samples implies reproducibility in the instrumental and environmental conditions during the data collection periods. The average ($\pm$ standard deviation) for the NIR and Raman data sets are $(7.21 \pm 3.51) \times 10^{-6}$ and $(1.72 \pm 0.82) \times 10^{-3}$ AU, respectively.

Analysis of the 100% lines was also used for the spectral quality of the NIR spectra. Each raw single-beam spectrum collected for a sample was ratioed to the other two spectra collected for that sample. The base-10 logarithm was computed and the root-mean-square (RMS) noise was calculated across each 100 cm^{-1} segment of the resulting 100% line. The RMS noise is computed relative to a second order polynomial fit of the 100% line data for each 100 cm^{-1} segment. Figure II-4 presents a set of representative 100% lines across the full NIR spectral range and the measured RMS noise levels for each 100 cm^{-1} segment are summarized in Table II-2. The results of the 100% lines are consistent with the results of the ratioed spectra. The lowest RMS noise levels correspond to where the water absorption is lowest. The average ($\pm$ standard deviation) for the NIR 100% lines data set is 3.13 ± 1.52 μAU.

Individual PLS Calibration Models

Separate PLS calibration models were built for each analyte with an optimized spectral range and number of latent variables or factors used in the model. The optimal spectral range and number of latent variables were determined by a grid search with a 10 cm^{-1} step and a window size ranging from 100 cm^{-1} to the full spectral range in

Table II-2. RMS noise level of 100% lines for NIR spectra.

NIR Spectral Range (cm^{-1})	4200-4300	4300-4400	4400-4500	4500-4600	4600-4700	4700-4800
NIR Noise Level (μAU)	41.1	6.85	3.13	3.32	4.79	13.2

increments of 10 cm^{-1} for both the NIR and Raman spectra with the lowest significant leave-one-out cross validation standard error of prediction (CVSEP). Figure II-5 shows the relationship between the number of factors and CVSEP. Standard error of calibration (SEC) and standard error of prediction (SEP) values were computed to quantify the performance of each model and the results are summarized in Table II-3.

For both the NIR and Raman models, the optimal spectral ranges for glucose, urea and lactate include the major spectral bands for each component. The SEC and SEP values summarized in Table II-3 indicate that NIR models for glucose, urea and lactate are smaller than the corresponding Raman models. Lower SEC and SEP values for the NIR models are consistent with lower RMS noise levels for the NIR data compared to the Raman data. For the NIR models, SEC and SEP values for glucose are higher than those for urea and lactate, which is an indication that the sensitivity over the combination spectral range is lowest for glucose.

Concentration correlation plots are presented in Figures II-6, II-7 and II-8 for both the NIR and Raman calibration models for glucose, urea and lactate, respectively. These plots show good calibration and prediction performance for these PLS models with predicted values closely distributed along the ideal correlation line.

Table II-4 summarizes results for the NIR and Raman PLS models computed over the full spectral ranges. For the NIR models, SEC and SEP values are much higher for the whole spectral range compared to the optimized range. For the Raman models, SEC and SEP values are similar to those for the optimized ranges. Optimized ranges for the NIR models are narrow compared to the whole range and represents an attempt to include the vital analytical absorption features while excluding noise. Optimized ranges for the Raman models are similar to the whole range, which is consistent with the fact that spectral quality is relatively constant across the spectrum as determined from the RMS noise levels detailed above. Both the optimized ranges for NIR spectra and Raman spectra are mainly discussed later.

Table II-3. Results for PLS calibration models computed over optimized spectral ranges for Raman and NIR spectra.

	Optimized Range (cm^{-1})		Number of Factors		SEC (mM)		SEP (mM)	
	Raman	NIR	Raman	NIR	Raman	NIR	Raman	NIR
Glucose	800-1500	4270-4440	5	5	0.36	0.26	0.54	0.27
Lactate	700-1550	4290-4500	5	5	0.53	0.11	0.46	0.10
Urea	900-1500	4500-4660	5	5	0.28	0.13	0.32	0.14

Table II-4. Results for PLS calibration models computed over the full spectral ranges for Raman and NIR spectra.

	Whole Range (cm^{-1})		Number of Factors		SEC (mM)		SEP (mM)	
	Raman	NIR	Raman	NIR	Raman	NIR	Raman	NIR
Glucose	700-1700	4200-4800	5	5	0.37	0.94	0.44	0.54
Lactate	700-1700	4200-4800	5	5	0.35	1.71	0.45	1.49
Urea	700-1700	4200-4800	5	5	0.26	1.23	0.34	0.75

Combined PLS Calibration Models

To judge the benefits of combining the NIR and Raman spectral data, PLS calibration models were generated by concatenating the NIR and Raman spectra for each sample. The analysis focused on results from the optimized spectral ranges.

Concatenation was accomplished by placing Raman intensity values along with the corresponding wavenumber values into consecutive column matrices and adding the NIR spectral values (signal and wavenumber) to follow the Raman values in the same matrix. The original point spacing for the Raman data was 0.3 cm^{-1}, and the point spacing for the NIR data was 1.9 cm^{-1}. To get similar point spacings for the two sets of spectra, the point spacing for the Raman spectra was changed to 2.1 cm^{-1} by taking every seventh data point from the original Raman data.

Results for the PLS models described above for the Raman data set were based on Raman spectra with a 0.3 cm^{-1} point spacing. The impact of modifying the point spacing from 0.3 to 2.1 cm^{-1} was tested by recomputing the PLS models for the full and optimized spectral ranges. The PLS models were built individually for Raman spectra with point spacings of 0.3 cm^{-1} and 2.1 cm^{-1}, and the SEC and SEP values for both data sets are identical, indicating no significant difference created in the modeling with this change in spectral point spacing.

Several different data transformations were used in attempts to match the NIR absorbance values with the Raman scattering values. The raw NIR absorption spectra have absorbance values on the order of 10^{-4} absorbance units. In contrast, the background subtracted Raman spectra have intensity values on the order of 10^{1} in arbitrary units. The follow transformations were investigated:

1) Raman + Raw NIR;
2) Raman + ($7 \times 10^{5} \times$ Raw NIR);
3) Normalized Raman + Normalized NIR;
4) Raman + NIR single-beam spectra;

5) Raman + (100×NIR single-beam spectra);
6) log (Raman sample / Raman blank) + NIR; and
7) log (Raman sample / Raman blank) + (10×NIR).

Raman spectra correspond to spectra after subtracting the background spectra as described above for the models summarized in Tables II-3 and II-4 and Figures II-6, II-7 and II-8. For the normalized Raman and NIR spectra, all values within a given spectrum were divided by the maximum value. For each transformation-concatenation method, PLS calibration models were generated over the full spectral ranges (4200-4800 cm^{-1} for NIR and 700-1700 cm^{-1} for Raman) and the optimized spectral ranges as determined from the individual Raman and NIR spectral analysis and tabulated in Tables II-14 and II-15. Figures II-9, II-10, II-11, II-12 and II-13 are examples of transformed concatenated spectra.

The first experiment generated PLS calibration models from sample spectra corresponding to the straight concatenation of the unmodified Raman and NIR spectral data. Results from this initial work are summarized in Table II-14 for the full spectral range and Table II-15 for the optimized range. SEC and SEP values for each analyte are similar to those obtained for PLS models based on the Raman spectra alone. Focusing on models based on the optimized spectral ranges denoted in Table II -5, SEC and SEP values for PLS models based on concatenated raw spectra are 0.36 and 0.54 mM for glucose, 0.53 and 0.46 mM for lactate and 0.28 and 0.32 mM for urea, respectively. Table II-5 summarizes SEC and SEP values for the individual models based on NIR and Raman alone (taken from Table II-3) and the concatenated models (taken from Table II-15).

The values in Table II-5 indicate that results from the concatenated models are identical to those from the Raman alone models. As indicated in Figure II-9, the magnitude of the raw Raman single is many orders of magnitude greater than that for the raw NIR spectra, raising the possibility that the combined model is dominated by the spectral variance associated with the Raman spectra. In this case, analyte information

Table II-5. PLS model results based on optimized spectral ranges with raw NIR and raw Raman spectra alone and concatenated raw Raman and raw NIR spectra.

	NIR Alone		Raman Alone		Raw Raman + Raw NIR	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.26	0.27	0.36	0.54	0.36	0.54
Lactate	0.11	0.10	0.53	0.46	0.53	0.46
Urea	0.13	0.14	0.28	0.32	0.28	0.32

embedded within the NIR spectra will be ignored and will not contribute to the resulting PLS model.

Before concatenation, the NIR spectra were amplified to better match the magnitude of the spectral signals. For this amplification, the raw NIR spectra were multiplied by a factor of 7×10^5. As shown in Figure II-10, the amplification puts the NIR and Raman spectra on the same magnitude. The PLS model based on concatenating raw Raman and $7 \times 10^5 \times$ raw NIR spectra resulted in the SEC and SEP values listed in Tables II-14 and II-15 for the full and optimized spectral ranges, respectively. Slight improvements are indicated compared to models generated from concatenating raw spectra of both types. Table II-6 summarizes the standard errors for the NIR and Raman alone models and the concatenated model based on raw Raman and amplified NIR spectra. From this tabulation of the SEC and SEP values, it is clear that the standard errors are reduced slightly compared to the raw Raman model, but are above those provided by the NIR alone models.

Another method to equalize the magnitude of the two spectral signals is to normalize the data relative to the maximum value within the data set (transformation method 3 specified above). In this case, all spectral values within each data set were normalized relative to the maximum signal of the spectra collected for the sixty samples. As illustrated in Figure II-11, this normalization process puts the two spectral types on the same order of magnitude. Results from the corresponding PLS models are presented in Tables II-14 and II-15. Focusing on models generated over the optimized spectral ranges, SEC and SEP values are essentially the same for glucose and urea and slightly improved for lactate, for the normalization method compared to the previous two transformation methods. All SEC and SEP values, however, are higher than the corresponding values obtained when using raw NIR spectra alone, as shown in Table II-7.

The remaining transformation methods were implemented as ways to put the spectral data into the same units. As noted above, the NIR data are in absorbance units

Table II-6. PLS model results based on optimized spectral ranges with raw NIR and raw Raman spectra alone and concatenated raw Raman and 7×10^5 amplified raw NIR spectra.

	NIR Alone		Raman Alone		Raw Raman + 7×10^5 NIR	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.26	0.27	0.36	0.54	0.32	0.43
Lactate	0.11	0.10	0.53	0.46	0.40	0.41
Urea	0.13	0.14	0.28	0.32	0.25	0.24

Table II-7. PLS model results based on optimized spectral ranges with raw NIR and raw Raman spectra alone and concatenated normalized Raman and normalized NIR spectra.

	NIR Alone		Raman Alone		Normalized Raman + Normalized NIR	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.26	0.27	0.36	0.54	0.34	0.49
Lactate	0.11	0.10	0.53	0.46	0.35	0.27
Urea	0.13	0.14	0.28	0.32	0.26	0.28

and the Raman data are in intensity units. Previous work has demonstrated the ability to create functioning PLS calibration models from analysis of single-beam NIR spectra.[14]

To see if the single-beam NIR spectra have an adverse effect on the quantitation from single-beam NIR spectra, PLS calibration models were compared where the input for one corresponded to the NIR absorbance spectra and the input for the other was the single-beam NIR spectra. Results from these models are summarized in Table II-8 and the resulting SEC and SEP values are similar. These findings imply no adverse impact on analytical measurements when using the single-beam NIR spectra.

For this reason, transformation method 4 explored the utility of using the Raman spectra and single-beam NIR spectra for of the sixty samples. As illustrated in Figure II-12, this concatenation process results in the Raman signal being several orders of magnitude larger than the NIR single-beam signal. Results from the corresponding PLS models are presented in Tables II-14 and II-15. Focusing on models generated over the optimized spectral ranges, SEC and SEP values are essentially the same for glucose, lactate and urea with the individual Raman model. All SEC and SEP values, however, are higher than the corresponding values obtained when using raw NIR spectra alone, as shown in Table II-9.

Noticing the magnitude of the NIR single-beam spectra is much lower than the Raman spectra, the transformation method 5 amplifies the NIR single-beam spectra by a multiplication factor of 100 before combining with the Raman spectra. As shown in Figure II-12, this amplification process puts the two types of spectra on the same order of magnitude. Results from the corresponding PLS models are presented in Tables II-14 and II-15. Focusing on models generated over the optimized spectral ranges, SEC and SEP values are similar for glucose, lactate and urea. All SEC and SEP values, however, are higher than the corresponding values obtained when using single-beam NIR spectra alone, as shown in Table II-10.

The other way to get the spectra into the same units is to use a log-ratio

Table II-8. PLS model results based on optimized spectral ranges with NIR absorbance spectra and NIR single-beam spectra.

	NIR Absorbance		NIR Single-beam	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.26	0.27	0.34	0.30
Lactate	0.11	0.10	0.14	0.16
Urea	0.13	0.14	0.13	0.12

Table II-9. PLS model results based on optimized spectral ranges with single-beam NIR and raw Raman spectra alone and concatenated Raman and single-beam NIR spectra.

	NIR single-beam Alone		Raman Alone		Raw Raman + single-beam NIR	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.34	0.30	0.36	0.54	0.36	0.54
Lactate	0.14	0.16	0.53	0.46	0.53	0.46
Urea	0.13	0.12	0.28	0.32	0.28	0.32

transformation of the Raman spectra, as indicated for the sixth and seventh transformations listed above. Inspection of the log-ratio Raman spectra reveals that the location and shape of the analyte peaks are essentially the same as those without the log-ratio treatment. Representative spectra are shown in Figure II-13. To see if the log-ratio transformation has an adverse effect on the quantitation from these Raman spectra, PLS calibration models were compared where the input for one corresponded to the raw (background subtracted) Raman spectra and the input for the other was the log-ratio transformed Raman spectra. Results from these models are summarized in Table II-11 and the resulting SEC and SEP values are similar. These findings imply no adverse impact on analytical measurements when implementing the log-ratio transformation with the Raman spectra.

Table II-12 compares results from PLS models for the Raman and NIR spectra alone along with the model using the sixth transformation where both the Raman and NIR spectra are concatenated in absorbance units. For the model from concatenated log-ratio spectra, the SEC and SEP values for the combination model are a little higher than the individual Raman model. This is due to the magnitude of the NIR spectra which is much lower than that of the transformed Raman spectra. The combined model is dominated by the Raman spectra with logarithm treatment. The addition of the NIR spectra with a low magnitude is like adding noise to the Raman spectra. This causes the higher SEC and SEP values of the combination model than the individual model.

As before, an attempt was made to bring the two data sets to the same order of magnitude. In this case, the NIR spectra were multiplied by a factor of 10. The corresponding PLS model results are presented in Table II-13. Small improvements are indicated compared to models generated from concatenating raw NIR spectra and log-ratio transformed Raman spectra. From this tabulation of the SEC and SEP values, it is clear that the standard errors are reduced slightly for lactate compared to the raw Raman model, and are similar to the log-transformed Raman model for glucose and urea,

Table II-10. PLS model results based on optimized spectral ranges with single-beam NIR and raw Raman spectra alone and concatenated Raman and 100×single-beam NIR spectra.

	NIR single-beam Alone		Raman Alone		Raw Raman + 100×single-beam NIR	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.34	0.30	0.36	0.54	0.34	0.54
Lactate	0.14	0.16	0.53	0.46	0.48	0.58
Urea	0.13	0.12	0.28	0.32	0.29	0.32

Table II-11. PLS model results based on optimized spectral ranges with raw Raman spectra and log-ratio Raman spectra.

	Raman		Log Raman	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.36	0.54	0.37	0.45
Lactate	0.53	0.46	0.37	0.34
Urea	0.28	0.32	0.26	0.28

Table II-12. PLS model results based on optimized spectral ranges with raw NIR and log-ratio Raman spectra alone and concatenated log-ratio Raman and NIR spectra.

	NIR Alone		Log Raman Alone		Log Raman + NIR	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.26	0.27	0.37	0.45	0.38	0.47
Lactate	0.11	0.10	0.37	0.34	0.40	0.44
Urea	0.13	0.14	0.26	0.28	0.30	0.33

Table II-13. PLS model results based on optimized spectral ranges with raw NIR and log-ratio Raman spectra alone and concatenated log-ratio Raman and 10 × NIR spectra.

	NIR Alone		Log Raman Alone		Log Raman + 10 × NIR	
Analyte	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)	SEC, (mM)	SEP, (mM)
Glucose	0.26	0.27	0.37	0.45	0.36	0.52
Lactate	0.11	0.10	0.37	0.34	0.36	0.26
Urea	0.13	0.14	0.26	0.28	0.28	0.34

but are above those provided by the NIR alone models.

Tables II-14 and II-15 summarize model results for glucose, lactate and urea for the different transformation methods when using the full and optimized spectral ranges, respectively. Among the different transformation methods, the best combined PLS model results are indicated for models based on raw Raman spectra combined with NIR absorbance spectra multiplied by a factor of 7×10^5. The corresponding concentration correlation plots are presented in Figure II-14. These plots show very good calibration and prediction ability of the combined PLS models with prediction values closely distributed along the ideal correlation line. However, from the SEC and SEP values of these models based on concatenated NIR-Raman spectra are only slightly better than those found for the Raman alone models and worse compared to the SEC and SEP values obtained for the NIR alone models.

Net Analyte Signal

Net analyte signal (NAS) is the orthogonal part of a spectrum relative to all other sources of spectral variance within a data set. For analytical purposes, the NAS is the unique portion of the analyte spectrum that can be used for selective quantitative measurements. In the case where a PLS calibration model is based on spectral features associated with the analyte of interest, the regression vector for the PLS model should be similar in structure compared to the corresponding NAS calibration vector generated for the same analyte in the same sample matrix. Figures II-15, II-16 and II-17 present superimposed NAS and PLS calibration vectors computed from Raman and NIR spectra for glucose, urea and lactate, respectively. For each analyte, the NAS and PLS vectors are similar, which indicates that information in the PLS model originates from the targeted analyte as opposed to spurious correlations within the spectral data set.[39]

The NAS calibration vector can be analyzed further to provide a measurement of the multivariate signal-to-noise ratio (SNR). In this analysis, the concentration

Table II-14. PLS model results for concatenated Raman and NIR spectra over full spectra ranges.

Transformation	Analyte								
	Glucose			Lactate			Urea		
	Rank	SEC	SEP	Rank	SEC	SEP	Rank	SEC	SEP
Raman + Raw NIR	5	0.37	0.44	5	0.35	0.45	5	0.26	0.34
Raman + ($7\times10^5\times$Raw NIR)	7	0.63	0.46	7	0.85	0.78	7	0.57	0.53
Normalized Raman + Normalized NIR	7	0.54	0.49	7	0.75	0.85	7	0.71	0.68
Raman + NIR single-beam spectra	6	0.37	0.44	6	0.35	0.45	6	0.26	0.34
Raman + (100$\times$NIR single-beam spectra)	6	0.37	0.43	6	0.45	0.53	6	0.27	0.33
log (Raman sample / Raman blank) + NIR	6	0.37	0.46	6	0.48	0.52	6	0.29	0.29
log (Raman sample / Raman blank) + (10$\times$NIR)	7	0.46	0.50	7	0.57	0.73	7	0.68	0.70

Table II-15. PLS model results for concatenated Raman and NIR spectra over optimized spectral ranges.

Transformation	Analyte								
	Glucose			Lactate			Urea		
	Rank	SEC	SEP	Rank	SEC	SEP	Rank	SEC	SEP
Raman + Raw NIR	5	0.36	0.54	5	0.53	0.46	5	0.28	0.32
Raman + ($7\times10^5\times$Raw NIR)	7	0.32	0.43	7	0.40	0.41	7	0.25	0.24
Normalized Raman + Normalized NIR	7	0.34	0.49	7	0.35	0.27	7	0.26	0.28
Raman + NIR single-beam spectra	6	0.36	0.54	6	0.53	0.46	6	0.28	0.32
Raman + ($100\times$NIR single-beam spectra)	6	0.34	0.54	6	0.48	0.58	6	0.29	0.32
log (Raman sample / Raman blank) + NIR	6	0.38	0.47	6	0.40	0.44	6	0.30	0.33
log (Raman sample / Raman blank) + ($10\times$NIR)	7	0.36	0.52	7	0.36	0.26	7	0.28	0.34

normalized length of the NAS vector represents the sensitivity of the measurement for the analyte. A measure of the noise can be obtained by projecting a set of blank spectra onto the NAS vector and computing the standard deviation of these projections. The ratio of the concentration normalized length of the NAS vector and the standard deviation of blank projections is analogous to a typical SNR calculation. Table II-16 lists the length of the concentration normalized NAS vector, noise, NAS-SNR along with the corresponding SEP values obtained from the PLS calibration models for glucose, urea and lactate. As expected, the NAS vector length is much larger for the Raman models compared to NIR models based on the magnitude and selectivity of the Raman emission bands. Correspondingly, the NAS noise for the Raman models is also much higher. The NAS-SNR for the Raman spectra is lower, however compared to the NIR model. The corresponding SEP values of the NIR spectra are lower than those of the Raman spectra, which is consistent with the rationalization that higher NAS-SNRs provide lower SEP values.

Similarly, Table II-17 summarizes NAS analysis values for Raman and NIR models computed over the full spectral ranges. For both types of spectra, NAS-SNR values decrease when the full range is considered, but the decrease is significantly more pronounced for the NIR data, especially for urea and lactate. Correspondingly, SEP values are worse as the NAS-SNR values decrease. The relationship between the NAS-SNR and these SEP values is shown in Figure II-19.

Table II-15 summarizes results for the best model based on concatenated spectra where Raman spectra are combined with multiplied NIR absorption spectra (multiplication factor = 7×10^5). A NAS vector analysis for this combined model was performed and the results are listed in Table II-18 and the resulting PLS regression vectors are presented in Figure II-18 for each analyte. Because more spectral information is included, the concentration normalized length of the corresponding NAS regression vectors is larger than that for either of the Raman or NIR models individually. The NAS

Table II-16. NAS calibration models for NIR spectra and Raman spectra for the optimized spectral range.

	NIR Models				Raman Models			
	NAS Length	NAS Noise	NAS SNR	SEP (mM)	NAS Length	NAS Noise	NAS SNR	SEP (mM)
Glucose	6.07×10^{-5}	8.01×10^{-6}	7.58	0.27	152.21	25.65	5.93	0.54
Lactate	1.34×10^{-4}	7.07×10^{-6}	18.95	0.10	128.32	35.08	3.65	0.46
Urea	6.62×10^{-5}	4.39×10^{-6}	15.06	0.14	222.76	32.16	6.92	0.32

Table II-17. NAS calibration models for NIR spectra and Raman spectra for the whole spectral range.

	NIR Models				Raman Models			
	NAS Length	NAS Noise	NAS SNR	SEP (mM)	NAS Length	NAS Noise	NAS SNR	SEP (mM)
Glucose	3.81×10^{-4}	7.79×10^{-5}	4.88	0.54	177.66	33.30	5.33	0.44
Lactate	2.83×10^{-4}	1.38×10^{-4}	2.04	1.49	134.71	38.64	3.48	0.45
Urea	9.61×10^{-4}	2.59×10^{-4}	3.70	0.75	240.30	41.05	5.85	0.34

Table II-18. NAS calibration models for concatenated Raman spectra and NIR spectra with a multiplication factor of 7×10^{5}.

	NAS Length	NAS Noise	NAS SNR	SEP (mM)
Glucose	173.51	25.88	6.70	0.43
Lactate	192.43	39.68	4.84	0.41
Urea	252.03	27.64	9.11	0.24

noise is similar to that of the Raman models. As a result, the NAS-SNR is higher than that of the Raman spectra but lower than that of the NIR spectra.

As noted above for the individual Raman and NIR models, the SEP is inversely related to the NAS-SNR. The green points plotted in Figure II-19 show this relationship for models based on concatenated spectra for each analyte. These points follow the same trend as those for the individual models.

Conclusions

The experimental results presented in this chapter demonstrate that the individual Raman spectra and NIR spectra can build valid PLS and NAS calibration models for glucose, urea and lactate in ternary mixtures. Several different strategies are evaluated for concatenating the Raman and NIR spectral data in order to assess the analytical utility of building multivariate calibration models based on combined Raman-NIR spectra. The best concatenation model was the second one which has the raw NIR spectra and the Raman spectra with amplification factor of 7×10^5. The SEC and SEP values of the combined spectra model fall in between the individual NIR model and individual Raman model.

The NAS SNR was examined for the two techniques separately and together. The SEP values have a decrease trend with the increase NAS-SNR. The SNR of combined model falls in between the SNR of the individual NIR model and individual Raman model. The lower the noise, the better prediction ability is the PLS model. Adding low NAS-SNR spectra to high NAS-SNR spectra reduces the prediction ability of the combined calibration model because of the additional noise.

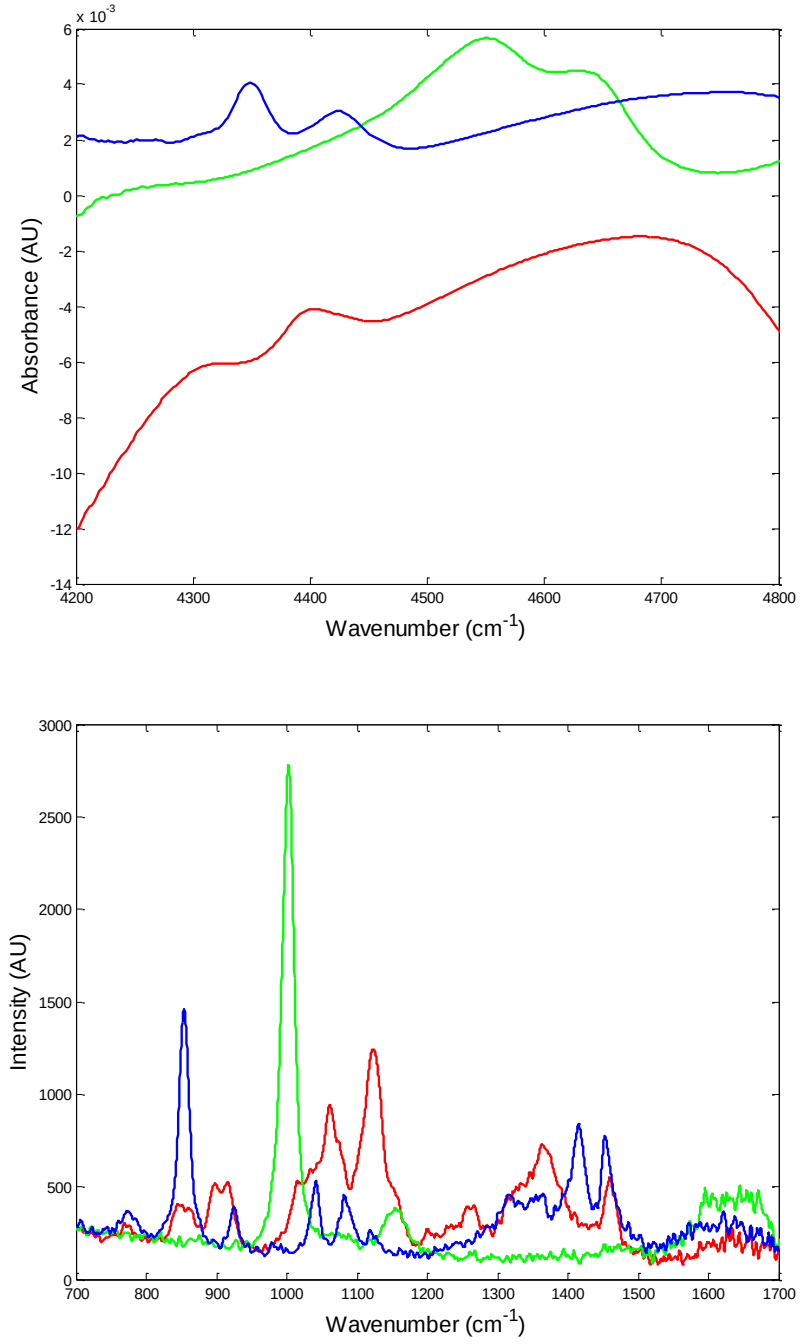

Figure II-1. NIR absorbance spectra (top) and Raman spectra (bottom) for individual 27 mM solutions of glucose (red), lactate (blue), and urea (green).

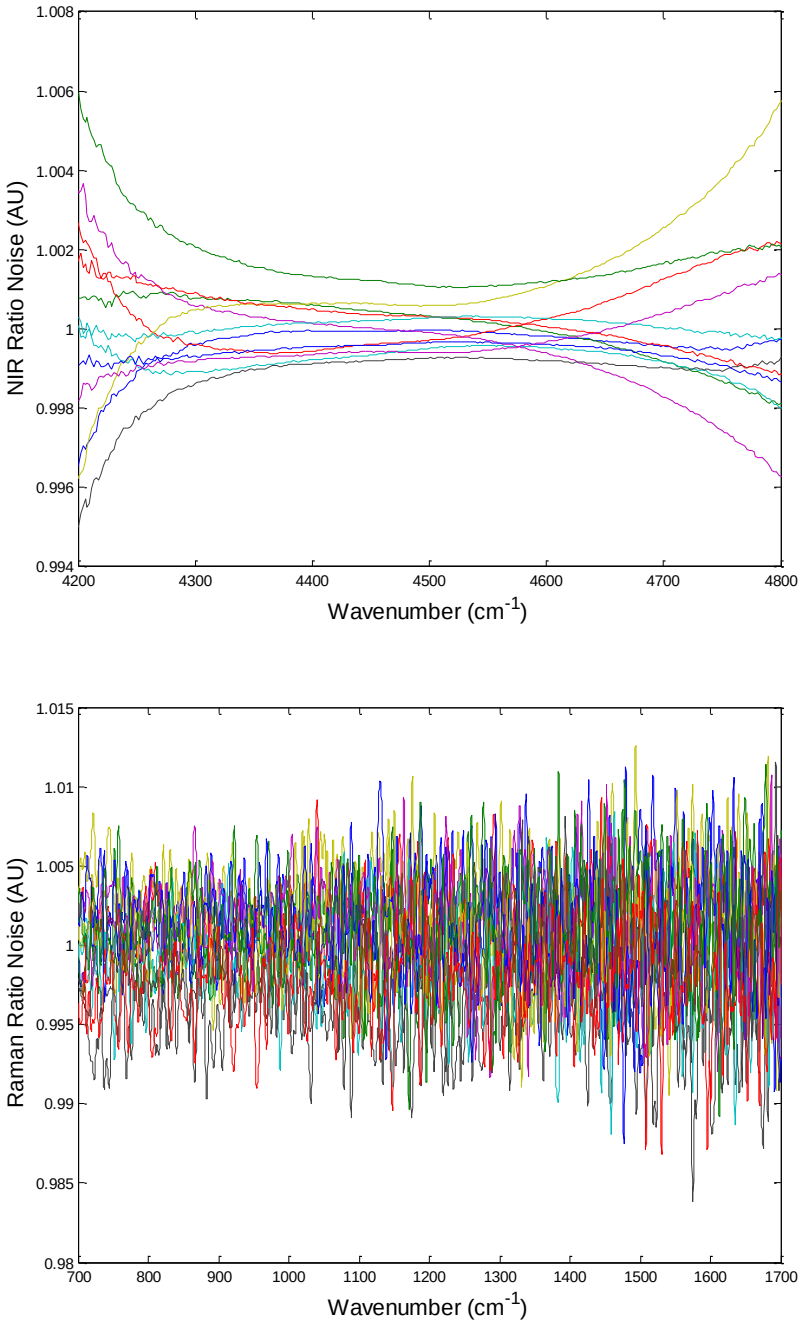

Figure II-2. Examples of ratioed spectra of NIR spectra (top) and Raman spectra (bottom).

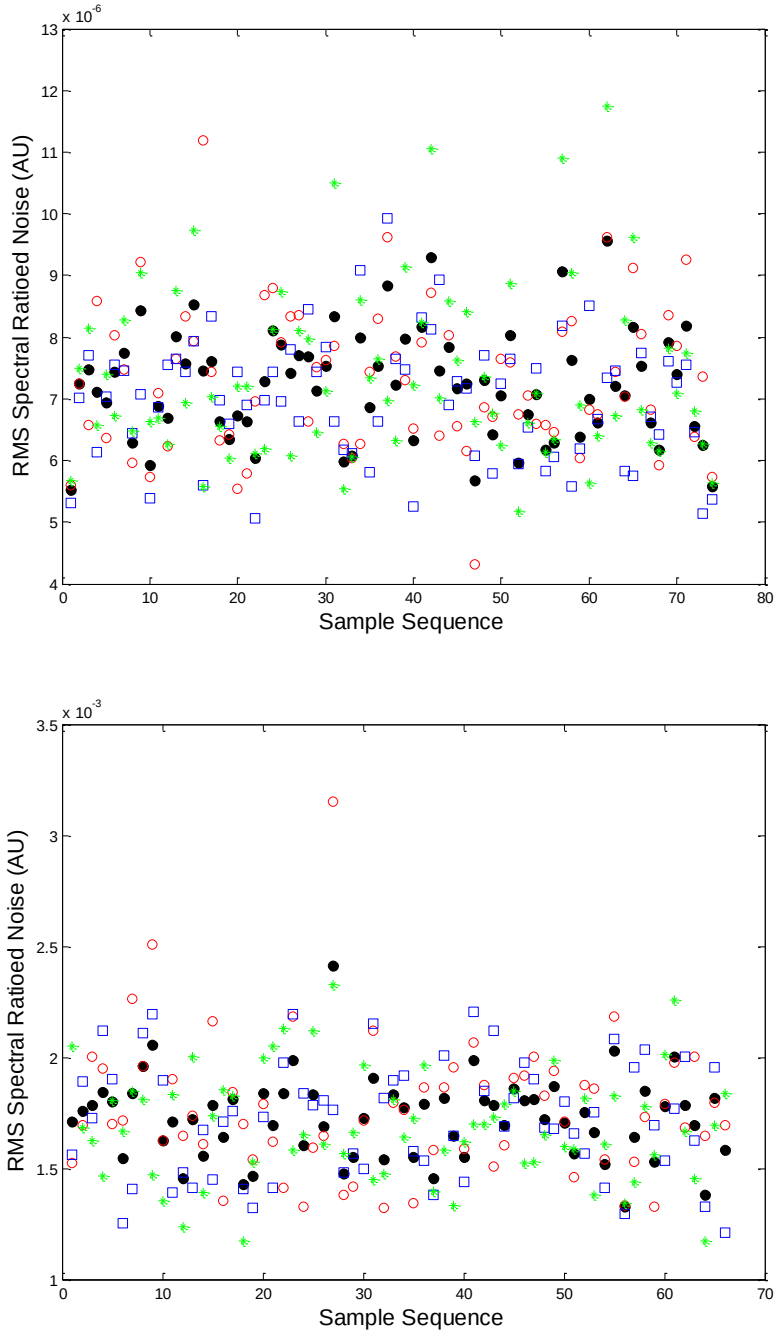

Figure II-3. Distribution of RMS noise on raioed spectra for NIR spectra (top) with a spectral range 4400-4500 cm-1and Raman spectra (bottom) with a spectral range 700-800 cm-1, showing values for the ratio of the first and second (red circles), second and third (blue squares), first and third (green stars) and the mean of the three (black circles) for each triplicate measurement.

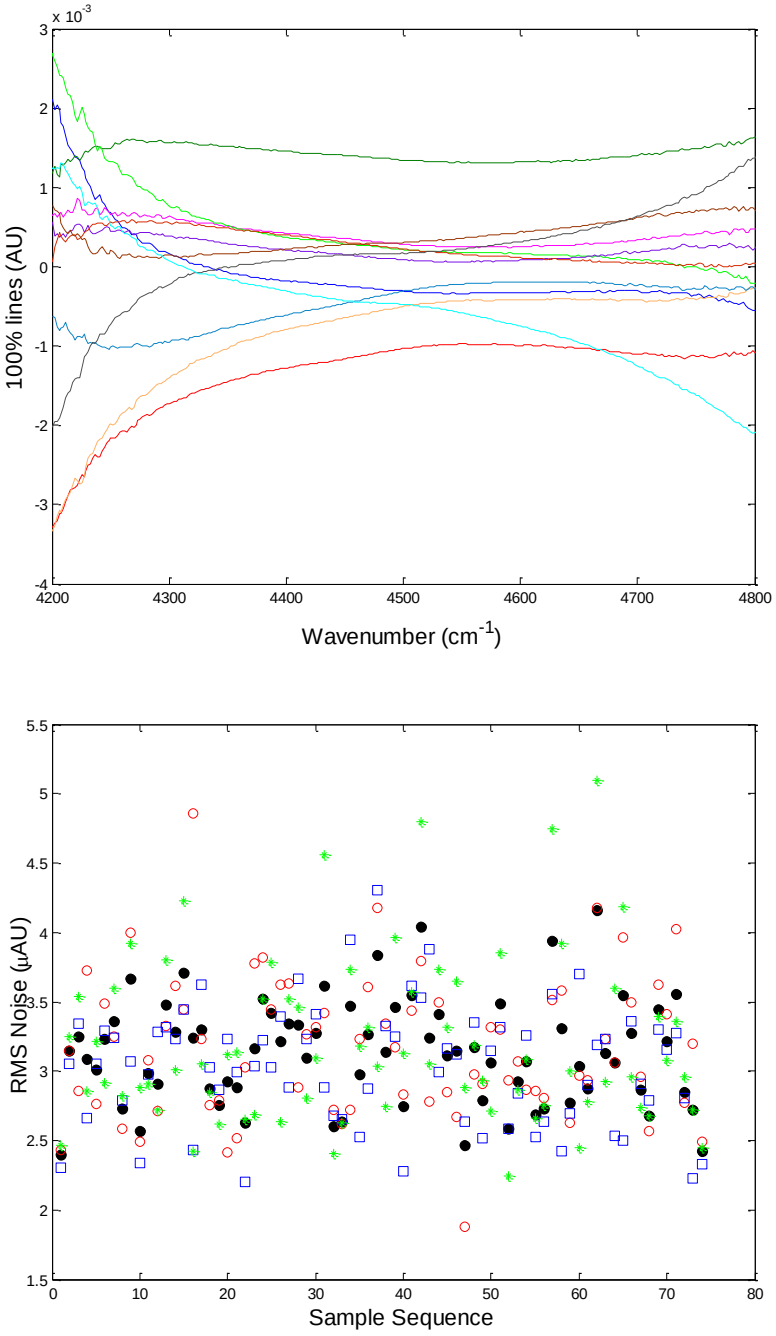

Figure II-4. Examples of 100% lines of NIR spectra (top) and distribution of RMS noise on 100% lines for NIR spectra (bottom) with a spectral range 4400-4500 cm-1, showing values for the ratio of the first and second (red circles), second and third (blue squares), first and third (green stars) and the mean of the three (black circles) for each triplicate measurement.

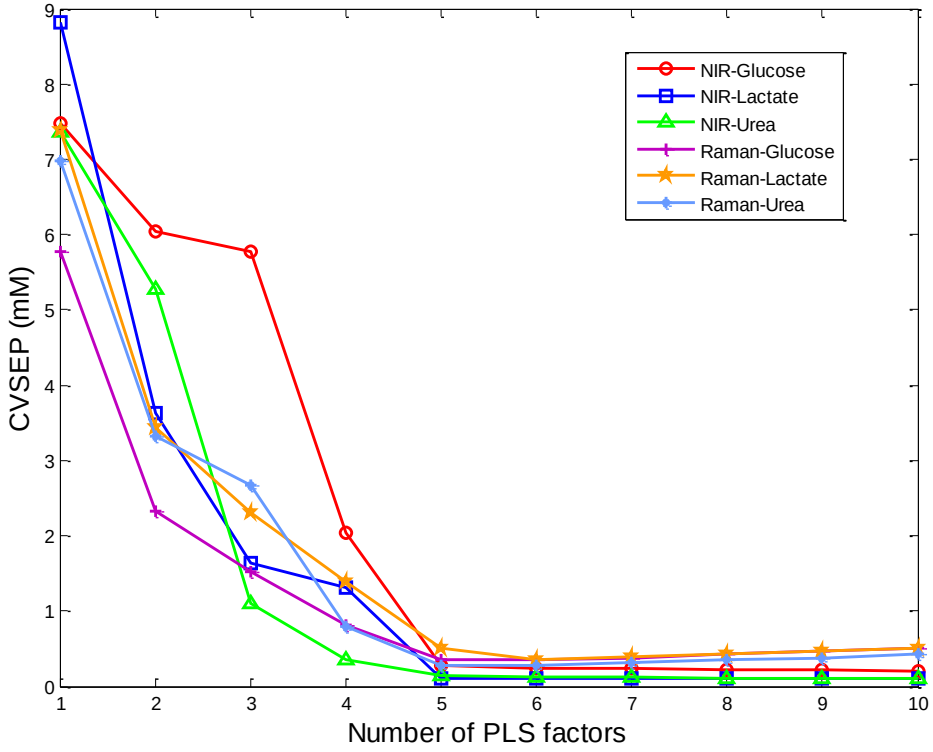

Figure II-5. Relationship between number of PLS factors and CVSEP for each individual analyte in both NIR spectra and Raman spectra.

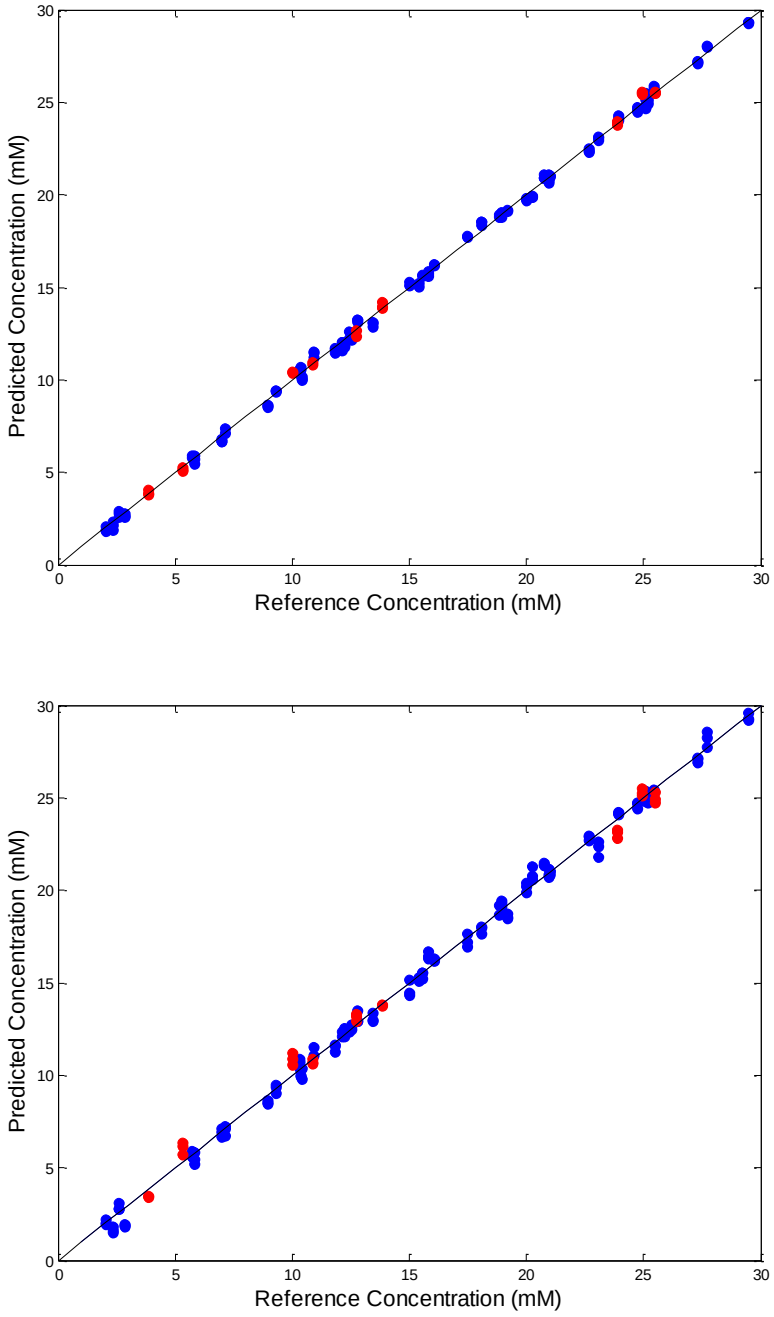

Figure II-6. Concentration correlation plots for glucose PLS calibration models from NIR spectra (top) and Raman spectra (bottom). Calibration data are represented by blue circles, and prediction data are represented by red circles. The black solid line is the ideal correlation.

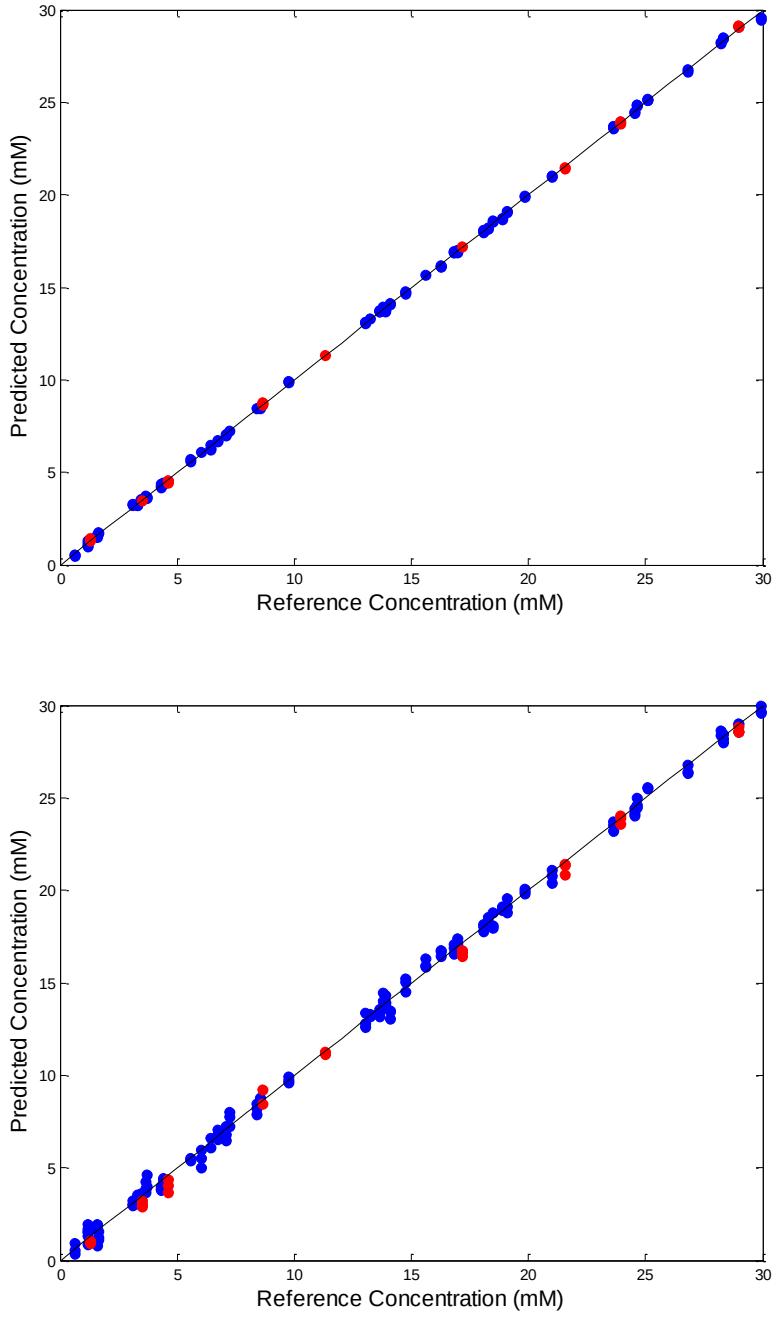

Figure II-7. Concentration correlation plots for lactate PLS calibration models from NIR spectra (top) and Raman spectra (bottom). Calibration data are represented by blue circles, and prediction data are represented by red circles. The black solid line is the ideal correlation.

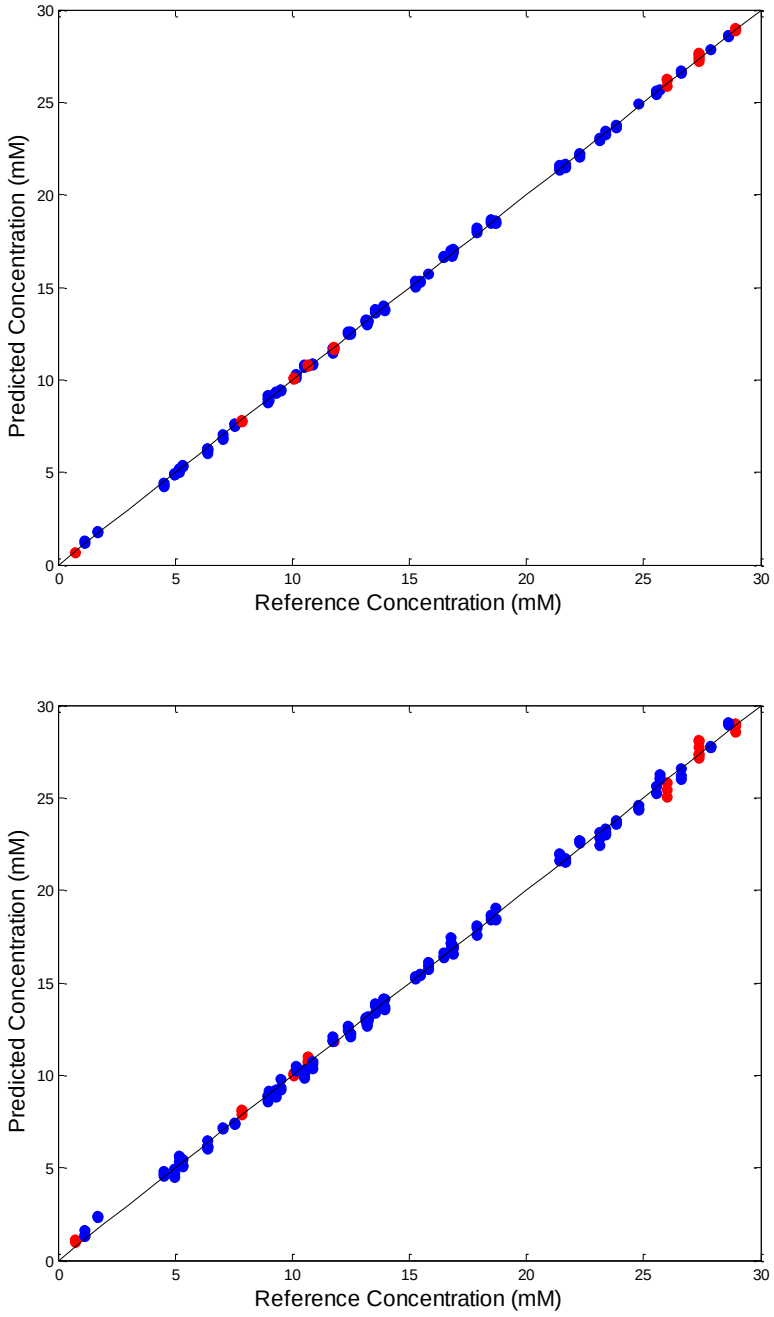

Figure II-8. Concentration correlation plots for urea PLS calibration models from NIR spectra (top) and Raman spectra (bottom). Calibration data are represented by blue circles, and prediction data are represented by red circles. The black solid line is the ideal correlation.

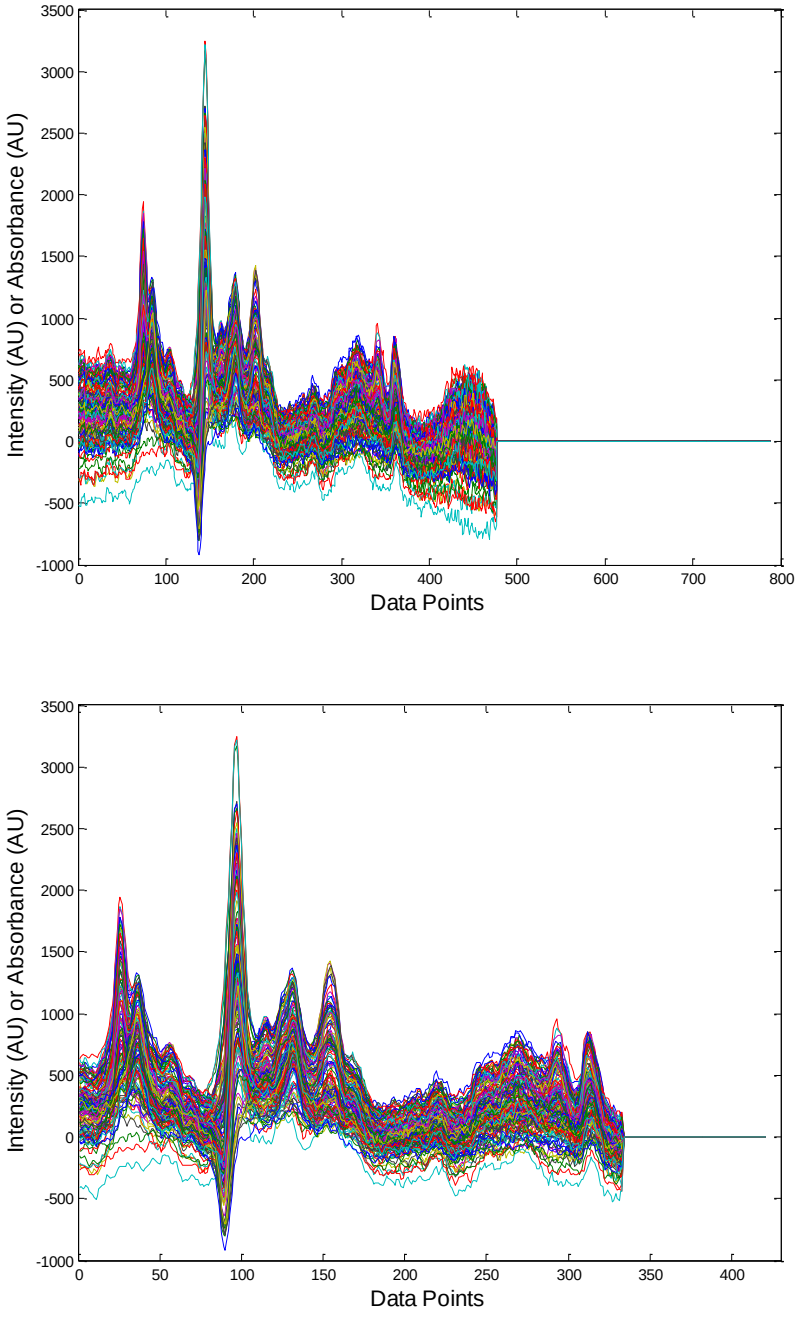

Figure II-9. Concatenated Raman and NIR spectra for the full spectral range (top) and optimized spectral range (bottom) for glucose using Transformation 1.

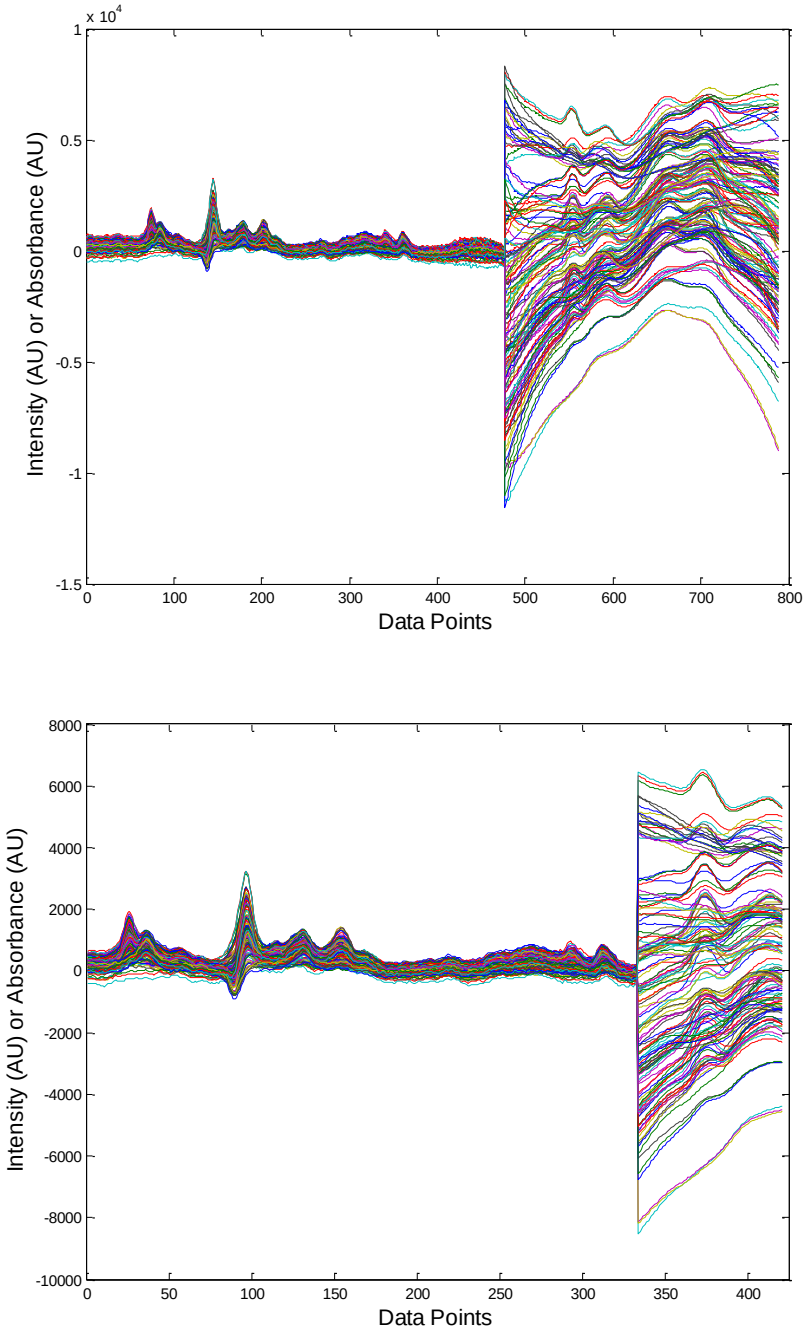

Figure II-10. Concatenated Raman and NIR spectra for the full spectral range (top) and optimized spectral range (bottom) for glucose using Transformation 2.

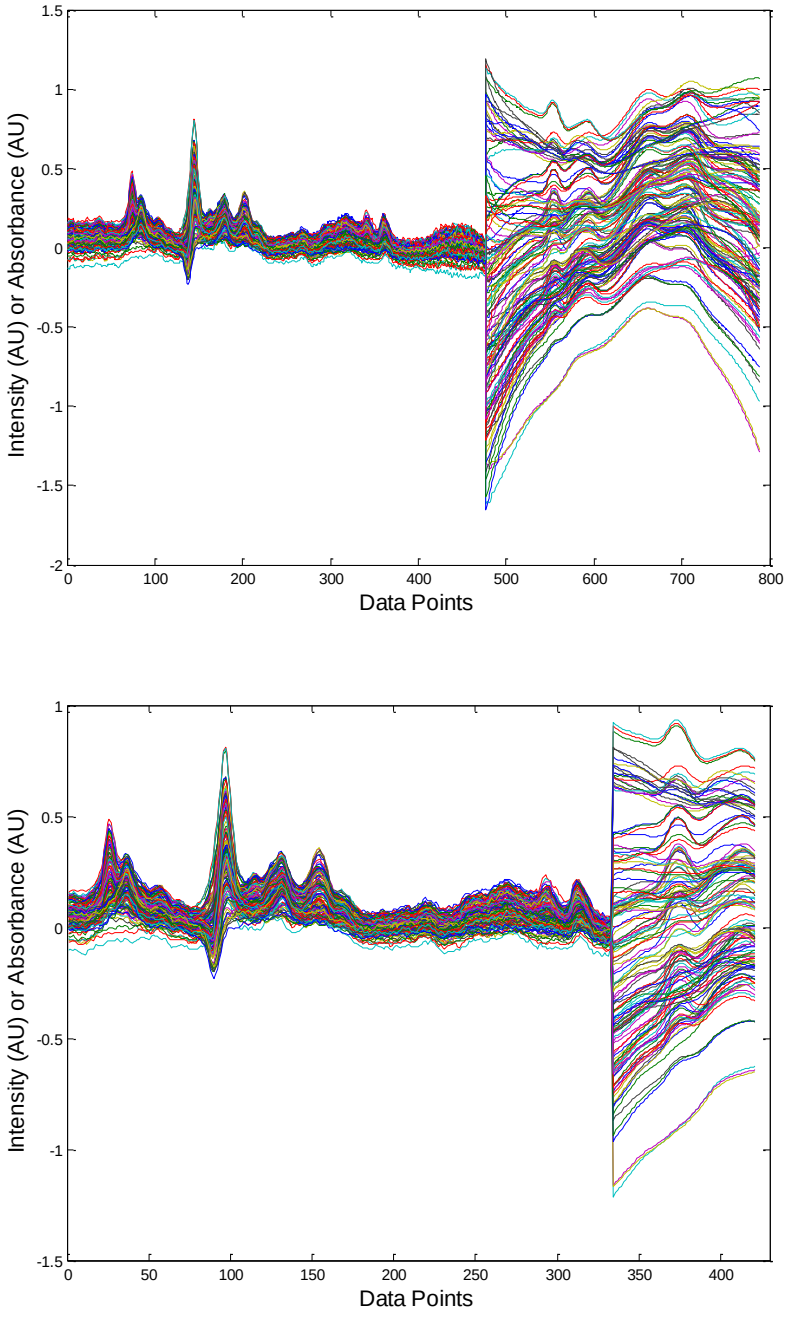

Figure II-11. Concatenated Raman and NIR spectra for the full spectral range (top) and optimized spectral range (bottom) for glucose using Transformation 3.

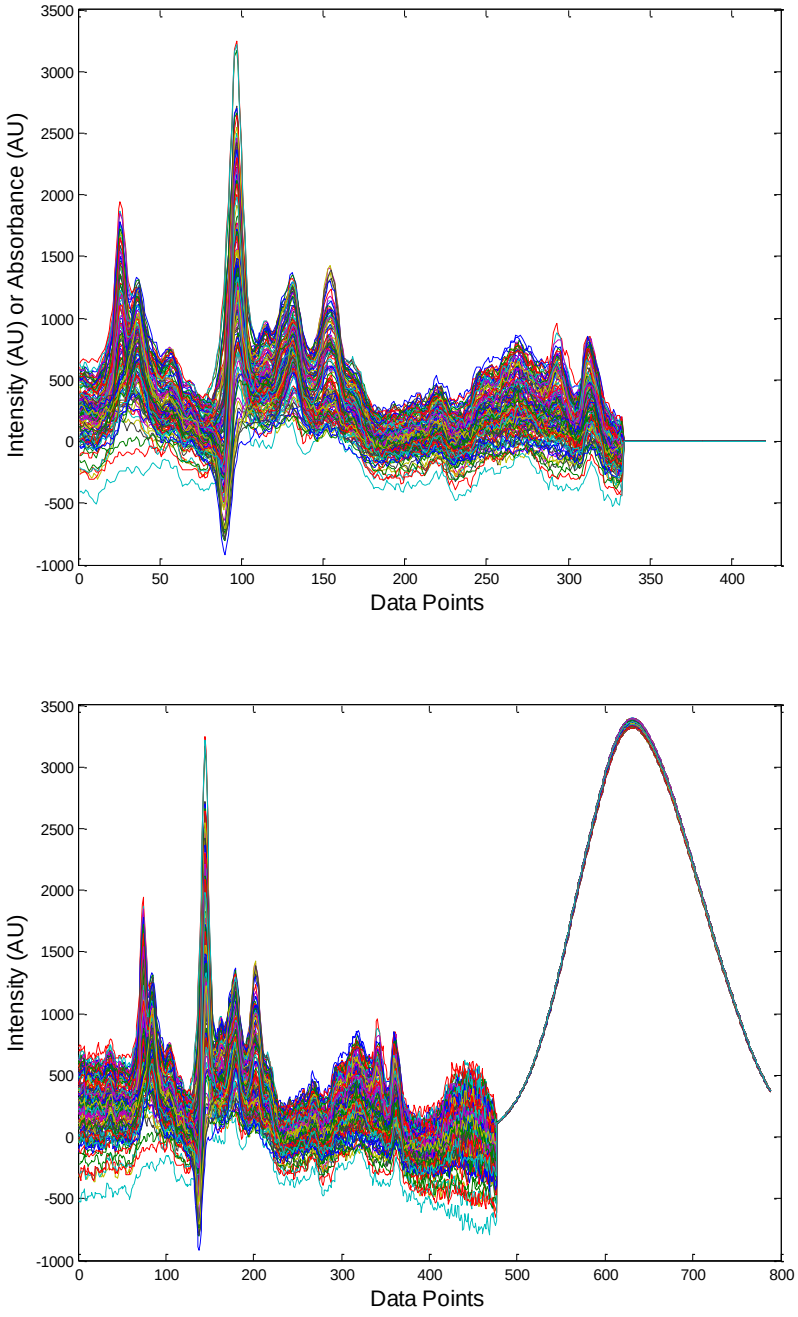

Figure II-12. Concatenated Raman and NIR spectra for optimized spectral range for glucose using Transformation 4 (top) and Transformation 5 (bottom).

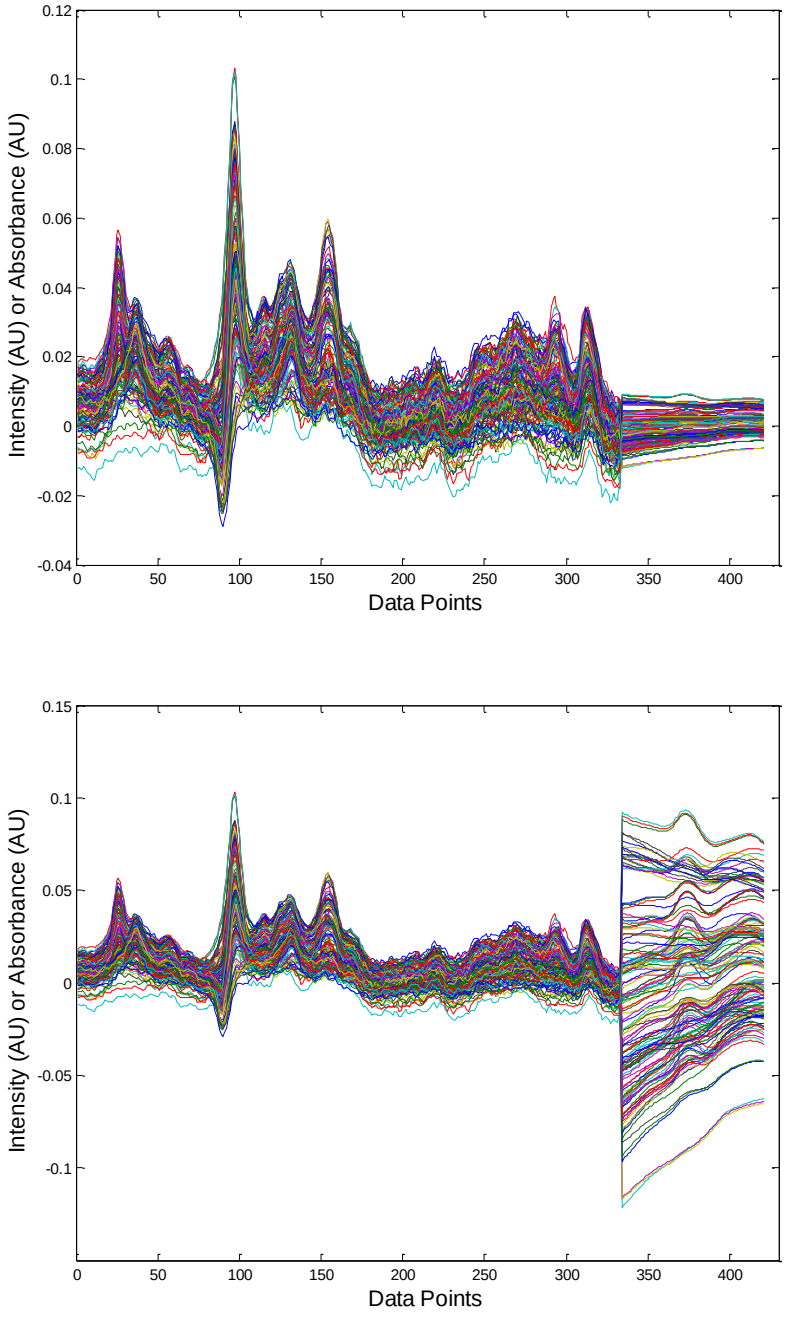

Figure II-13. Concatenated Raman and NIR spectra for optimized spectral range for glucose using Transformation 6 (top) and Transformation 7(bottom).

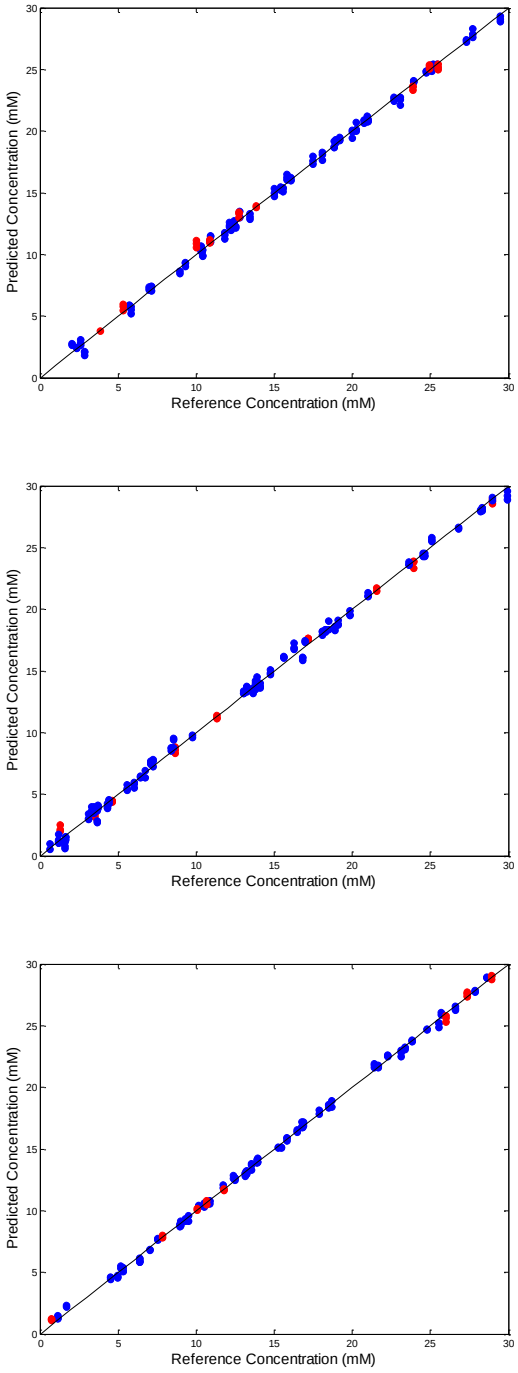

Figure II-14. Concentration correlation plots for combined Raman and NIR spectra PLS calibration models for glucose (top), lactate (middle) and urea (bottom). Calibration data are represented by blue circles, and prediction data are represented by red circles. The black solid line is the ideal correlation.

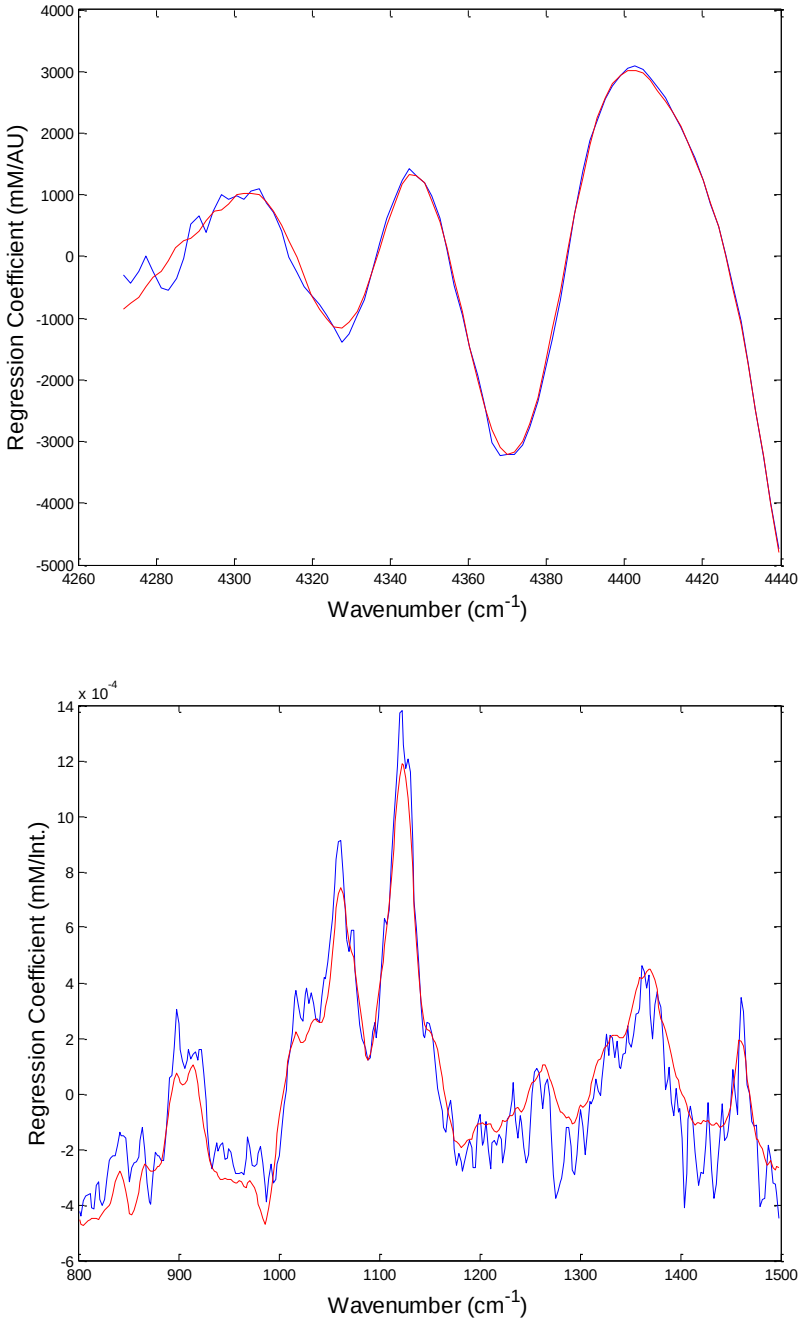

Figure II-15. Comparison of PLS (red) and NAS (blue) calibration vectors for glucose from NIR spectra (top) and Raman spectra (bottom).

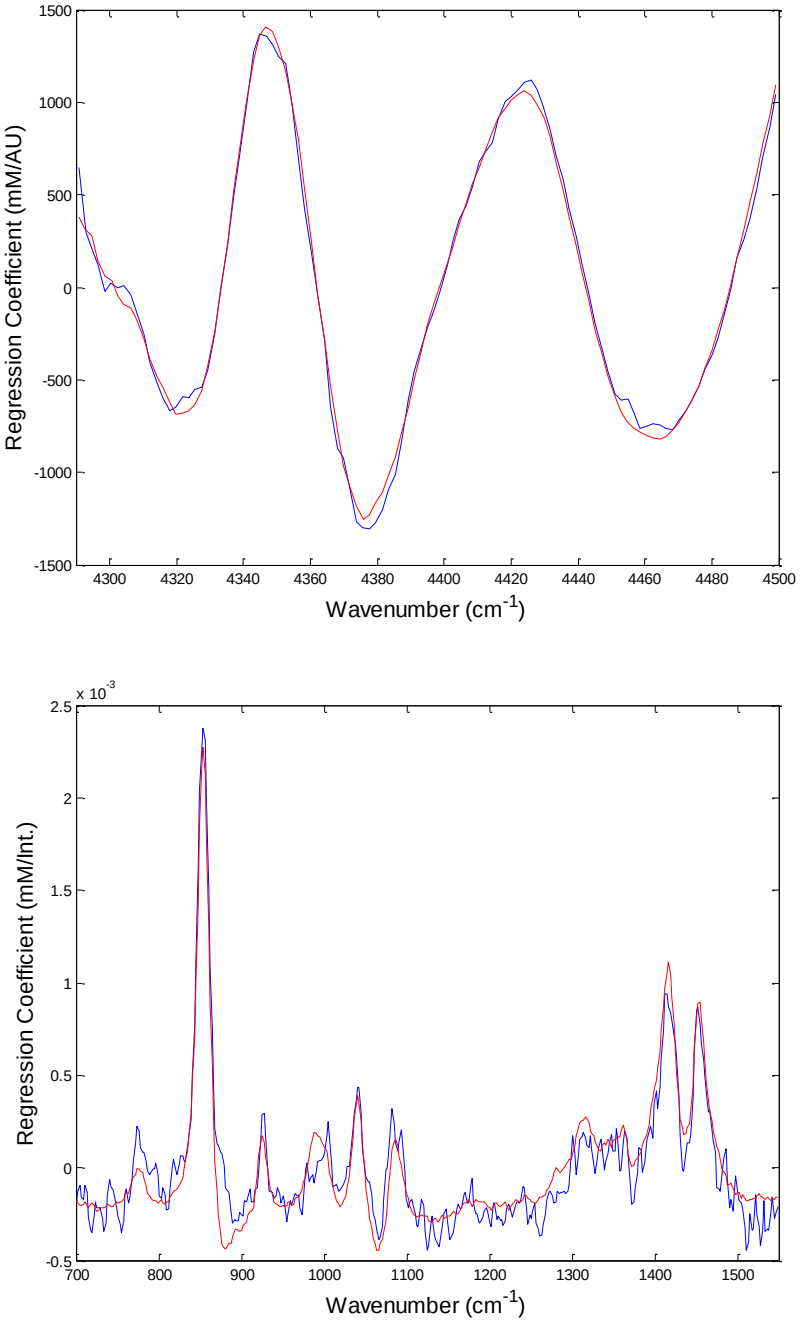

Figure II-16. Comparison of PLS (red) and NAS (blue) calibration vectors for lactate from NIR spectra (top) and Raman spectra (bottom).

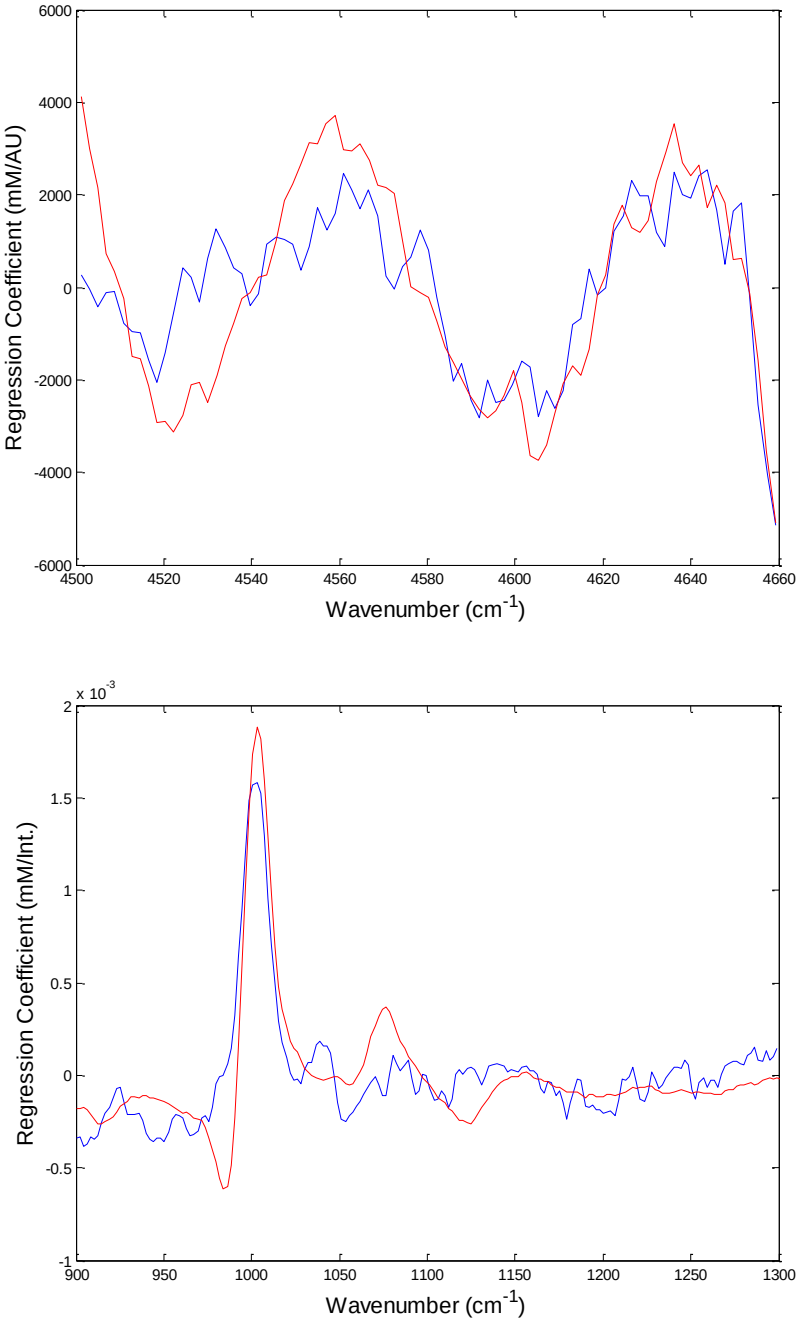

Figure II-17. Comparison of PLS (red) and NAS (blue) calibration vectors for urea from NIR spectra (top) and Raman spectra (bottom).

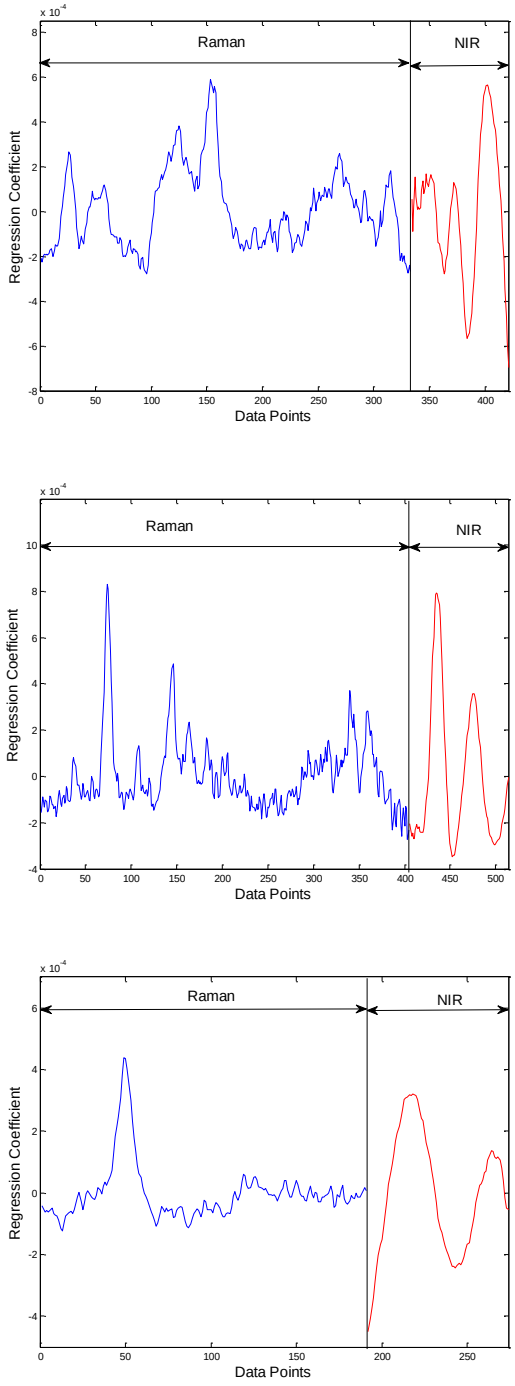

Figure II-18. PLS calibration vectors for combined Raman spectra (blue) and NIR spectra (red) for glucose (top), lactate (middle) and urea (bottom) in Transformation 2.

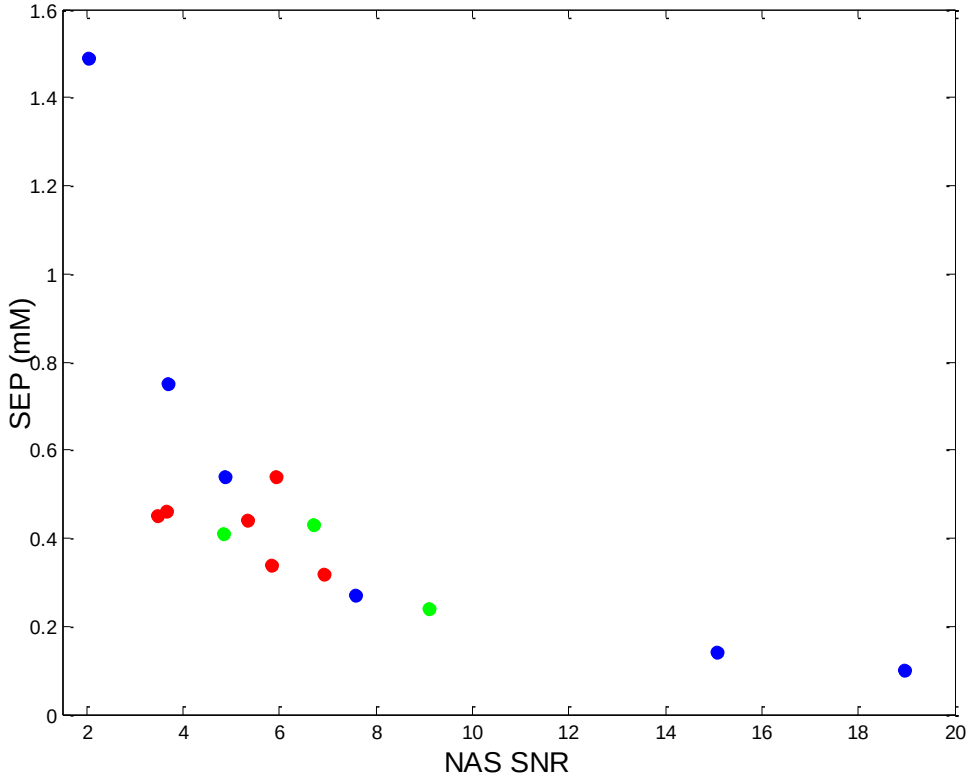

Figure II-19. Relationship between NAS-SNR and SEP values where blue dots correspond to values determined from individual NIR spectra, red dots correspond to models generated from individual Raman spectra and green dots are derived from models from concatenated NIR and Raman models with Transformation 2.

CHAPTER III

COMPARISON AND COMBINATION OF SIMULATED NEAR-INFRARED AND RAMAN SPECTRA FOR PLS AND NAS QUANTITATION OF GLUCOSE, UREA AND LACTATE

Introduction

In Chapter II, Raman and NIR spectra collected in the laboratory for a set of sixty ternary mixtures were used to assess the utility of combining these types of spectral data to improve the prediction ability of PLS calibration models for the measurements of glucose, urea, and lactate. Results indicate that improvement can be realized compared to individual PLS models generated from Raman spectra alone, but the predication ability of the combined Raman-NIR models is inferior to models based on NIR spectra alone. Analysis of the multivariate SNR provides some insight into the cause of these findings. A fundamental relationship is described between the NAS-SNR and the SEP of the resulting calibration models. Furthermore, analysis of the NAS-SNR from the experimental Raman and NIR spectra indicates that combining spectra of high SNR (NIR spectra in Chapter II) with spectra of low SNR (Raman spectra in Chapter II) will produce models that are inferior to those generated by the high SNR spectra alone.

In this chapter, Raman spectra and NIR spectra are simulated as a means to verify and expand the findings reported in Chapter II. The ability to control noise levels within the simulated spectra provides a valuable tool to examine the impact of noise on the final calibration models.

Experimental Section

Molar absorptivity values for glucose, urea and lactate were taken from Amerov.[35] Absorbance values were computed according to Equation III-1 which takes into account absorptions by the analyte and water. As described in detail elsewhere, the effect of water displacement within the optical volume of the measurement is relevant

owing to the non-zero absorption properties of water.[35] All NIR spectra were created assuming an optical path length of 1 mm and a point spacing of 1.9 cm^{-1}. Simulated absorbance spectra for 1 mM concentrations of pure solutions of glucose, urea and lactate are presented in Figure III-1.

$$A_{sample} = \sum A_{solute} - \sum A_{water\ displacement} \qquad \text{Equation III-1}$$

Simulated Raman spectra were generated from the pure component Raman spectra described in the previous chapter. Pure emission spectra for glucose, urea, and lactate were generated by subtracting averaged buffer spectra from averaged spectra collected for the 27 mM pure solutions of each analyte, respectively. The baseline was further corrected to zero by subtracting a first polynomial fitted baseline function computed over the 700-1700 cm^{-1} spectral range. Noise was removed from the resulting spectral baseline by assigning zero for wavenumber values where no emission bands were evident. Point spacing for the simulated Raman spectra was set to 2 cm^{-1}. The resulting pure component spectrum for each analyte was divided by 27 to normalize to 1 mM and this concentration normalized spectrum was used to simulate mixtures assuming a linear relationship between emission intensity and molar concentration. The final concentration normalized pure component simulated Raman spectra for glucose, urea and lactate are presented in Figure III-1.

Noise was generated as Gaussian distributed random numbers with the MATLAB function ‘randn’ with a distribution centered at zero. Magnitude of the noise was adjusted by multiply these random numbers with an appropriate factor. The noise was added to the simulated spectra to form spectra with a given level of noise.

For simulated Raman and NIR spectra, the ternary mixture spectra were generated with the same concentrations of glucose, urea and lactate represented in the sixty mixtures described in Chapter II. For the NIR simulated spectra, a multiplication factor of

7.6×10^5 was applied. This multiplication factor is similar to the value used in Transformation 2 described in Chapter II and corresponds to the value required to match the lengths of the concentration normalized NAS calibration vectors for the Raman and NIR spectral data. Two levels of spectral noise were investigated. Initially, a noise level of ± 150 arbitrary units was used to match the relative noise levels observed in the Raman spectra described in Chapter II. Subsequently, the impact of noise was evaluated by doubling the noise level to ± 300. Examples of simulated NIR and Raman spectra with $\pm$ 150 noise levels for 27 mM pure component solutions of glucose, lactate and urea are presented in Figure III-2. NAS-based multivariate parameters of NAS vector length, NAS-noise, and NAS-SNR are listed below in Tables III-1 and III-2 for each analyte and for each noise level.

Results and Discussion

PLS models for the simulated Raman and NIR spectra were built as described in Chapter II. Results for the individual PLS models based on simulated Raman and NIR spectra at the two noise levels are tabulated in Table III-3. The optimized ranges are the same as the optimized ranges reported in Chapter II.

The number of factors is 3 for each of the PLS model reported in Table III-3 and the SEC and SEP values are similar for models based on Raman spectra and NIR spectra when the same noise level is used for each. As expected, doubling the noise level increases the standard errors by a factor of two. Both PLS and NAS calibration vectors are presented in Figures III-3, III-4 and III-5 for the individual analytes. The NAS and PLS regression vectors are similar in structure and magnitude, as expected when the PLS models originate from the target analyte.

PLS models were established for combined Raman and NIR spectra, in an analogous fashion to that described in Chapter II and the results are presented in Table III-4. Again, the spectral ranges used for the concatenated spectra are the optimized

Table III-1. NAS calibration parameters for simulated NIR and Raman spectra over the optimized spectral ranges with a noise level of ±150.

	NIR Models				Raman Models			
	NAS Length	NAS Noise	NAS SNR	SEP (mM)	NAS Length	NAS Noise	NAS SNR	SEP (mM)
Glucose	192.34	32.37	5.94	0.27	191.97	31.24	6.14	0.19
Lactate	189.15	39.35	4.80	0.30	155.90	33.53	4.64	0.31
Urea	292.69	41.53	7.04	0.25	233.89	34.17	6.84	0.18

Table III-2. NAS calibration parameters for simulated NIR and Raman spectra over the optimized spectral ranges with noise levels of ±300.

	NIR Models				Raman Models			
	NAS Length	NAS Noise	NAS SNR	SEP (mM)	NAS Length	NAS Noise	NAS SNR	SEP (mM)
Glucose	192.46	88.48	2.17	0.57	193.15	65.95	2.92	0.51
Lactate	193.44	96.45	2.00	0.62	160.74	70.92	2.26	0.56
Urea	327.26	90.61	3.61	0.39	239.25	76.44	3.12	0.47

Table III-3. Results from individual PLS calibration models based on simulated Raman and NIR spectra.

	Noise level	Rank		SEC (mM)		SEP (mM)	
		Raman	NIR	Raman	NIR	Raman	NIR
Glucose	±150	3	3	0.24	0.26	0.19	0.27
	±300	3	3	0.52	0.52	0.51	0.57
Lactate	±150	3	3	0.32	0.29	0.31	0.30
	±300	3	3	0.66	0.57	0.56	0.62
Urea	±150	3	3	0.20	0.16	0.18	0.25
	±300	3	3	0.41	0.25	0.47	0.39

Table III-4. Results of PLS calibration models based on concatenated simulated Raman and NIR spectra.

Methods[a]	Glucose			Lactate			Urea		
	Rank	SEC	SEP	Rank	SEC	SEP	Rank	SEC	SEP
(1)	3	0.15	0.13	3	0.21	0.20	3	0.09	0.08
(2)	3	0.30	0.36	3	0.42	0.43	3	0.20	0.25
(3)	3	0.25	0.23	3	0.37	0.37	3	0.16	0.15
(4)	3	0.23	0.27	3	0.31	0.27	3	0.15	0.19

[a] (1) Raman and NIR spectra with noise levels of ±150, (2) Raman and NIR spectra with noise levels of ±300, (3) Raman spectra with noise levels of ±150 and NIR spectra with noise level of ±300, and (4) Raman spectra with noise levels of ±300 and NIR spectra with noise levels of ±150.

ranges for Raman and NIR spectra described in Chapter II. Four different combinations of the Raman and NIR spectra are indicated in Table III-4 where: 1) the combined Raman and NIR spectra have noise levels of ± 150; 2) the Raman and NIR spectra have noise levels of ± 300; 3) the Raman spectra have a noise level of ± 150 and the NIR spectra have a noise level of ± 300; and 4) and the Raman spectra have a noise level of ± 300 and the NIR spectra have a noise level of ± 150.

When the Raman and NIR spectra have the same noise of ±150, the SEC and SEP values in the combined model are lower than the corresponding SEC and SEP values in the individual models with the same noise level. The SEC and SEP in these models are dictated by the magnitude of the spectral noise. This finding indicates that if the SNR of the Raman spectra and NIR spectra are at the same level, concatenating individual spectra improves the prediction ability of the multivariate models relative to models based on individual Raman or NIR spectra.

For the method combining Raman and NIR spectra with the same noise level of ± 300, the SEC and SEP values in the concatenated model are lower than the corresponding SEC and SEP values in the individual model with the same noise level. This result further indicates that if the SNR of the Raman spectra and NIR spectra are at the same level, the prediction ability of the combined Raman and NIR model is better than the individual Raman or NIR model. PLS regression vectors from concatenated models are presented in Figure III-6 for each analyte. Because of the multiplication factor of 7.6×10^5 applied to the NIR spectra, magnitudes of the Raman and NIR portions of the total regression vector are the same. The concentration correlation plots for the concatenated Raman and NIR PLS models with noise levels of ± 150 for each of component are presented in Figure III-7. These plots show good calibration and prediction ability of the PLS models with prediction values closely distributed along the ideal unity line.

When the noise level for the Raman spectra is lower (± 150 compared to ± 300 for the NIR spectra) or when the noise level for the NIR spectra is lower (± 150 compared to

±300 for the Raman spectra), the model performance is between the two extremes. As indicated in Table III-4, SEC and SEP values are between the cases where the noise is ± 150 and ±300 for both spectral types within the modeled data set. This finding implies that adding low SNR spectra to high SNR spectra has an adverse effect on the prediction ability of the combined model and is consistent with the findings determined in Chapter II.

Table III-5 summarizes the NAS vector length, NAS noise and NAS-SNR and the corresponding SEP values from the PLS models of the four different combined models. The NAS-SNR is plotted relative to the SEP in Figure III-8. As was found in Chapter II, a decreasing trend is observed for SEP with an increase in NAS-SNR. For the combined Raman-NIR models with different noise levels, the SEC and SEP values are higher or similar with the individual PLS model with the lower SEC and SEP values but lower than the individual PLS model with high SEC and SEP.

Conclusions

Models generated from simulated spectra show that the prediction ability of the PLS model improves when the NAS-SNR of the two spectra are similar. This finding means that both the magnitude of the NAS vector length and the noise level should be similar for the combined data sets. If there is a significant difference in the NAS vector magnitude, the model would be dominated by the spectra with higher magnitude. The SEC and SEP values fall between the individual models if the two spectra have different NAS-SNR values. This finding indicates that adding low NAS-SNR spectra to high NAS-SNR spectra will make the prediction ability of the combined model worse than that generated by the high NAS-SNR spectral data. Under conditions of constant noise, higher NAS magnitudes correspond to greater sensitivity and improved analytical measurements. Likewise, for constant sensitivity (NAS magnitude) lower NAS-noise provides superior response.

Table III-5. NAS calibration parameters for concatenated simulated NIR and Raman spectra for the optimized spectral range with different noise levels.

Methods[a]	Analyte	NAS length	NAS noise	NAS SNR	SEP
(1)	Glucose	329.09	33.60	9.79	0.13
	Lactate	250.08	31.74	7.87	0.20
	Urea	459.51	35.99	12.76	0.08
(2)	Glucose	337.13	73.77	4.57	0.36
	Lactate	255.84	76.04	3.36	0.43
	Urea	459.22	84.21	5.45	0.25
(3)	Glucose	332.07	55.16	6.02	0.23
	Lactate	251.09	56.29	4.46	0.37
	Urea	458.29	52.25	8.77	0.15
(4)	Glucose	330.78	62.52	5.29	0.27
	Lactate	257.64	47.97	5.37	0.27
	Urea	457.69	57.86	7.91	0.19

[a] (1) Raman and NIR spectra with noise levels of ± 150, (2) Raman and NIR spectra with noise levels of ± 300, (3) Raman spectra with noise levels of ± 150 and NIR spectra with noise level of ± 300, and (4) Raman spectra with noise levels of ± 300 and NIR spectra with noise levels of ± 150.

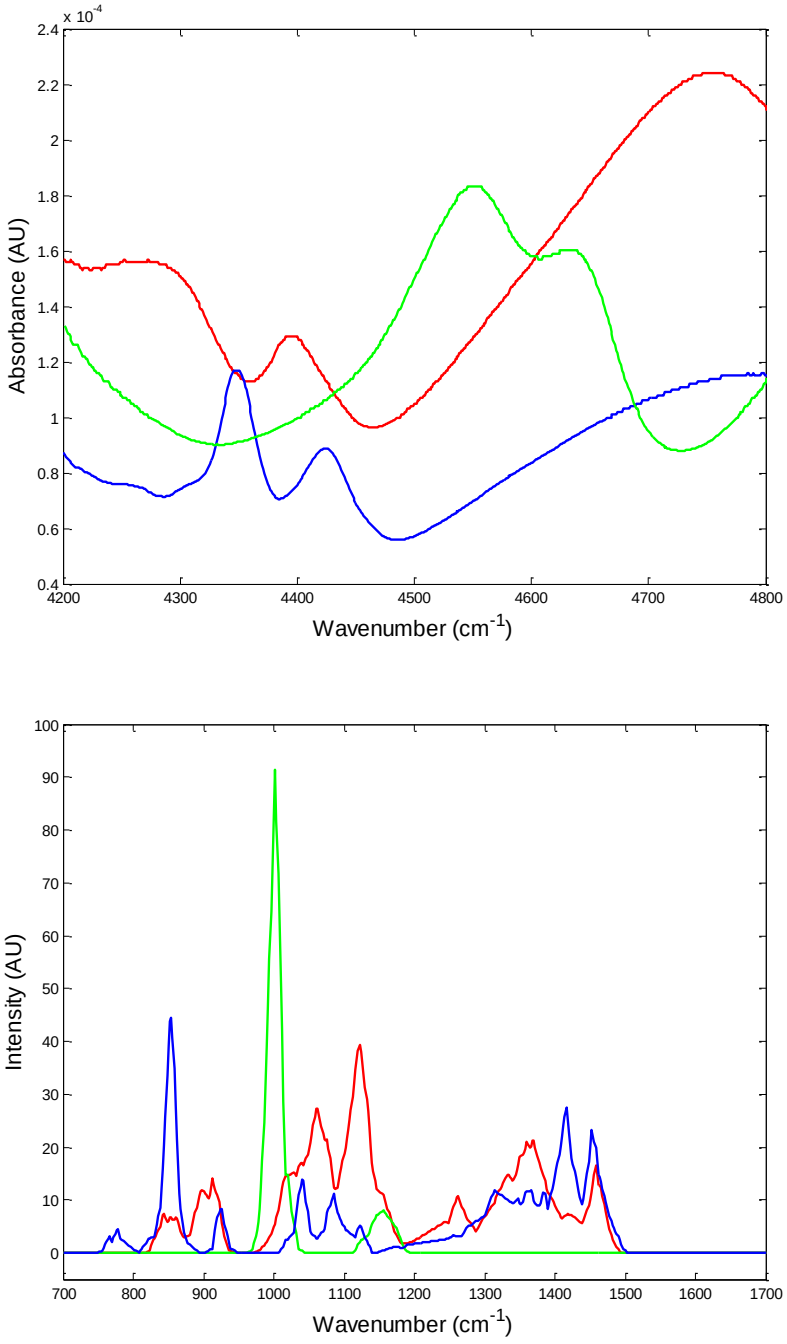

Figure III-1. Simulated pure component NIR absorbance spectra (top) and Raman spectra (bottom) normalized to 1mM for glucose (red), lactate (blue), and urea (green).

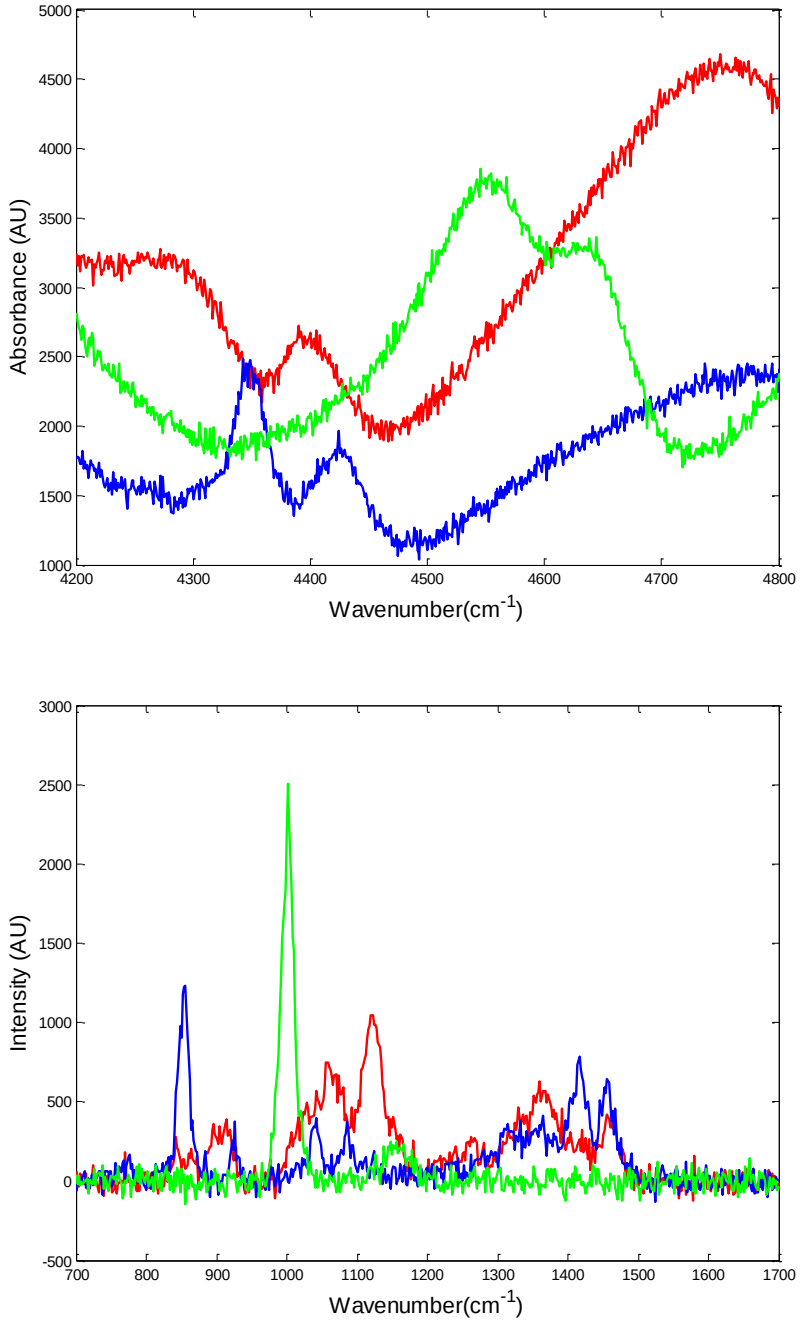

Figure III-2. Simulated pure component NIR absorbance spectra (top) and Raman spectra (bottom) with noise levels of ± 150 for 27 mM concentrations of glucose (red), lactate (blue), and urea (green) with noise.

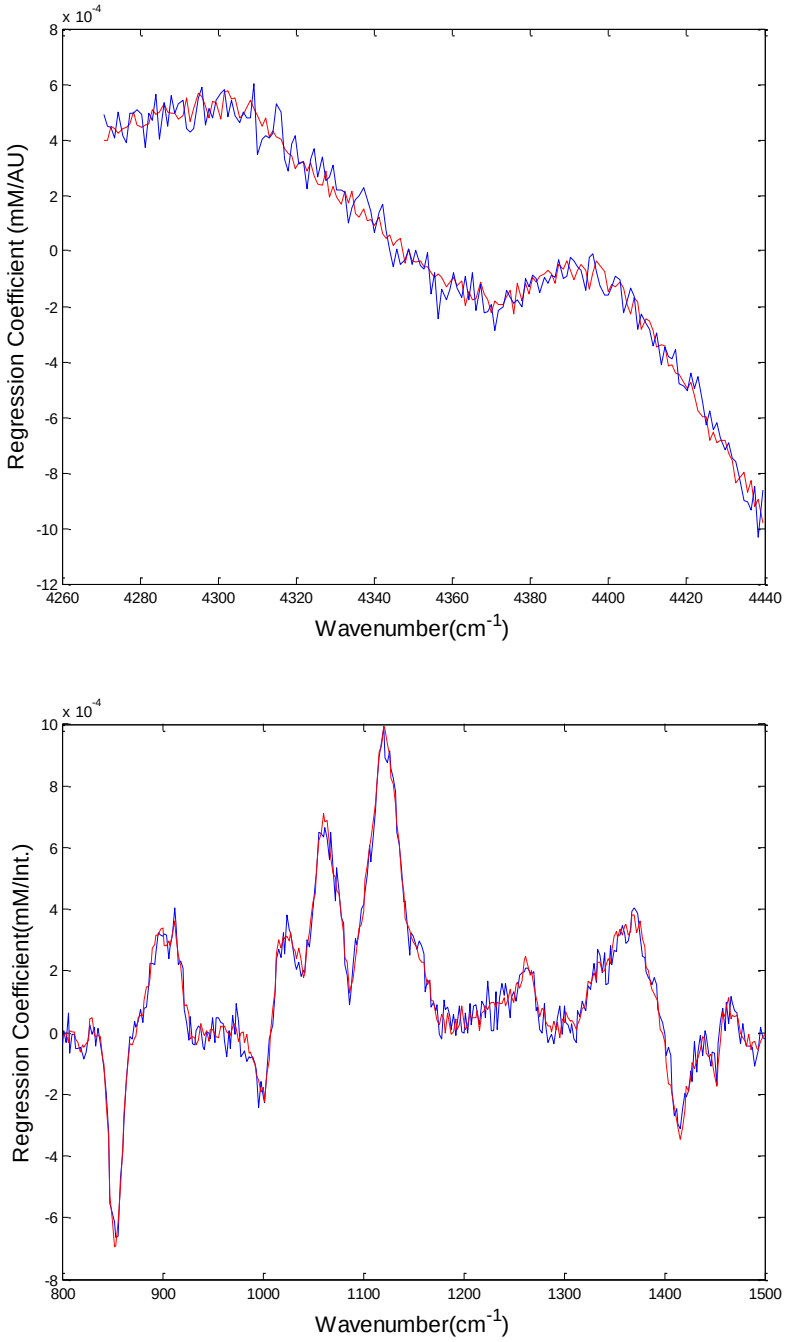

Figure III-3. Comparison of PLS (red) and NAS (blue) calibration vectors for glucose from NIR spectra (top) and Raman spectra (bottom).

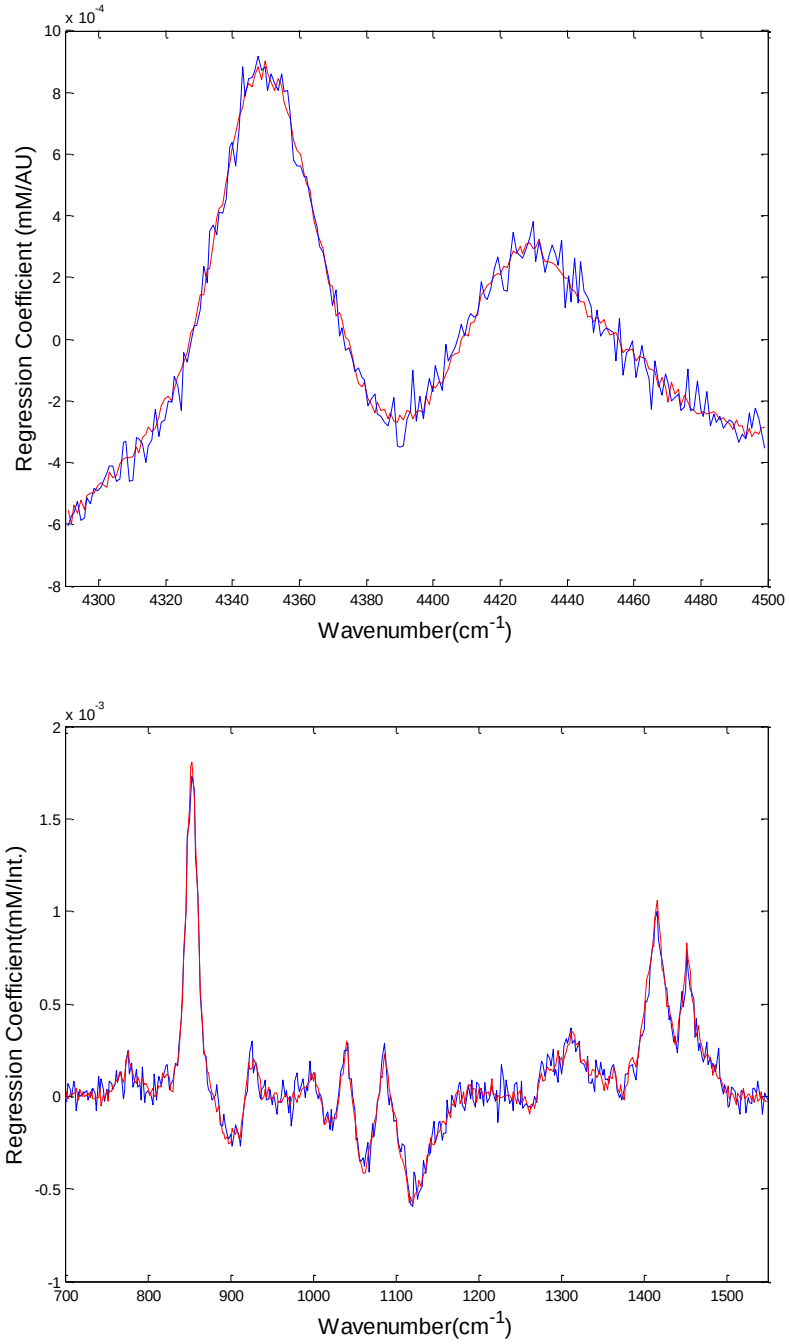

Figure III-4. Comparison of PLS (red) and NAS (blue) calibration vectors for lactate from NIR spectra (top) and Raman spectra (bottom).

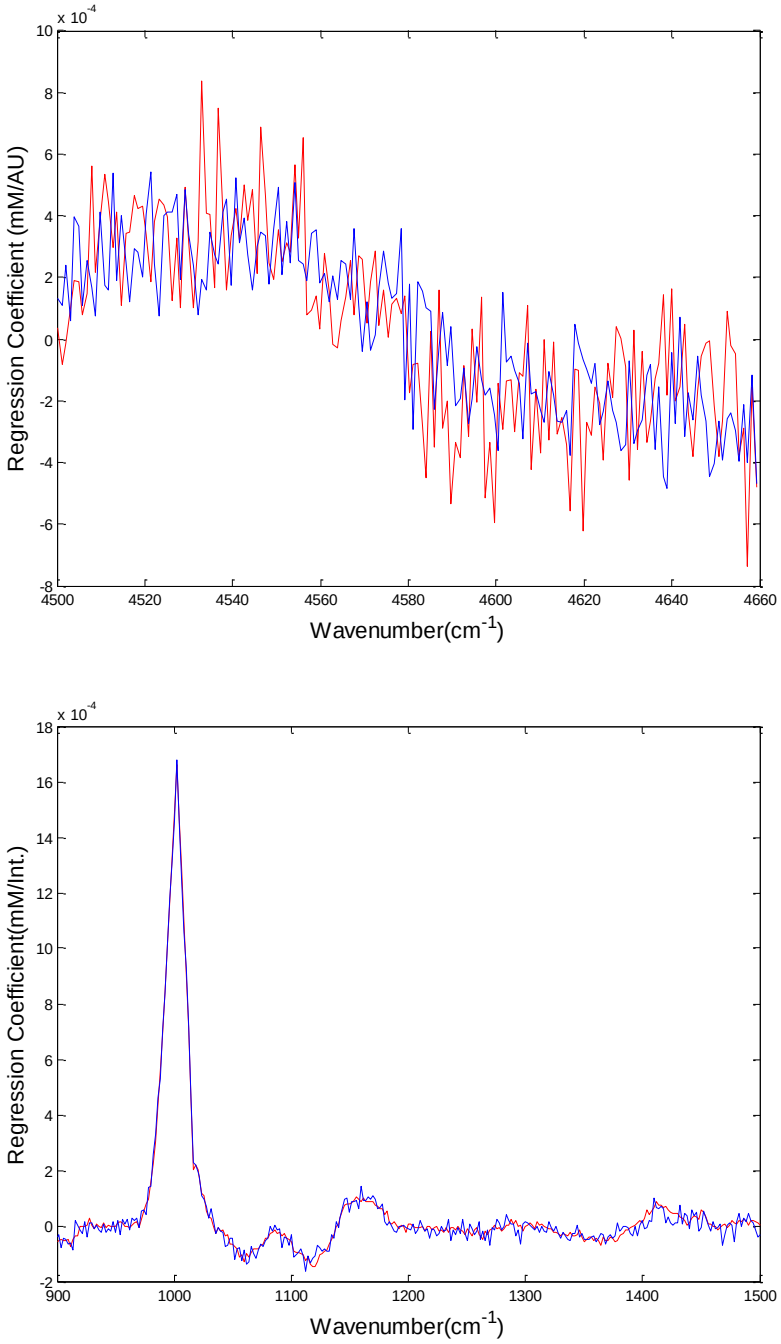

Figure III-5. Comparison of PLS (red) and NAS (blue) calibration vectors for urea from NIR spectra (top) and Raman spectra (bottom).

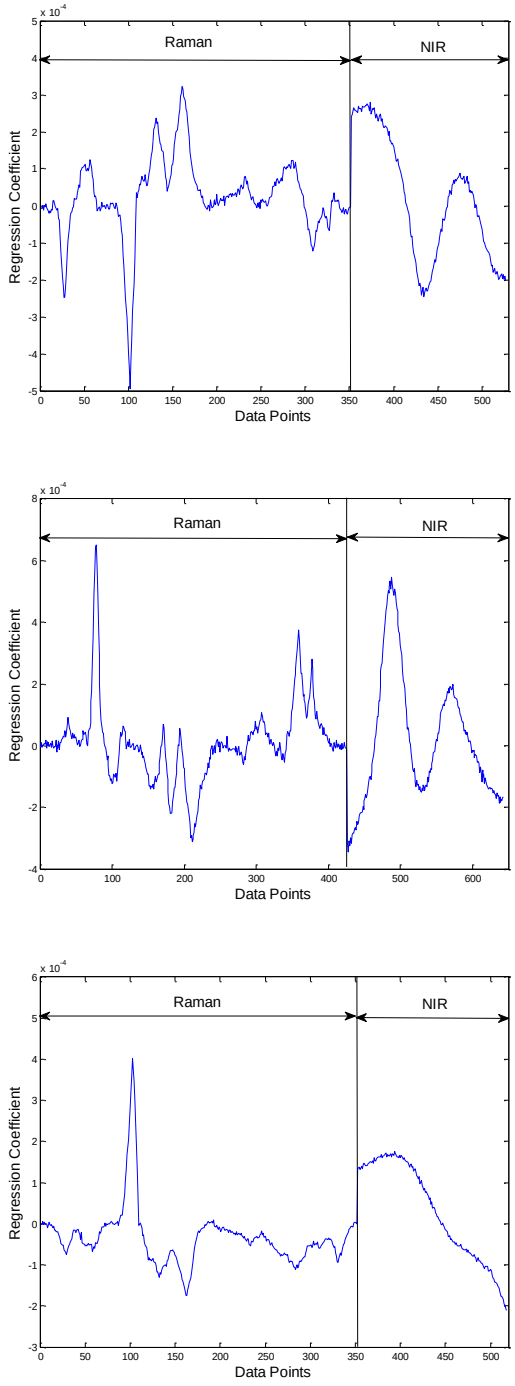

Figure III-6. PLS calibration vectors for combined Raman spectra and NIR spectra for glucose (top), lactate (middle) and urea (bottom).

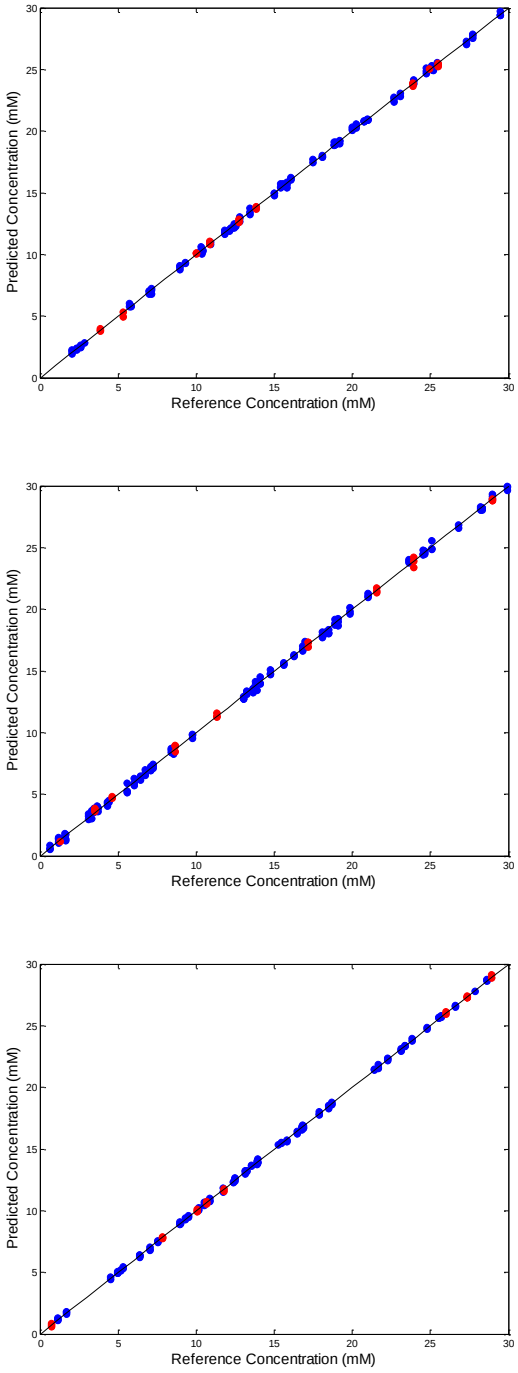

Figure III-7. Concentration correlation plots for combined Raman and NIR spectra PLS calibration models for glucose (top), lactate (middle) and urea (bottom). Calibration data are represented by blue circles, and prediction data are represented by red circles. The black solid line is the ideal correlation.

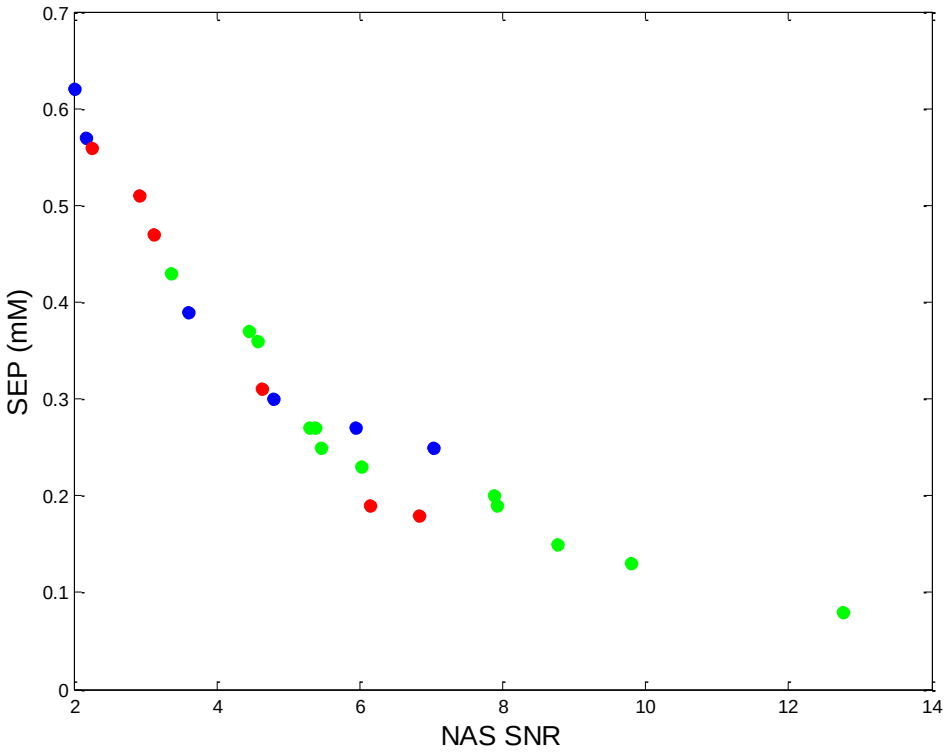

Figure III-8. Relationship between NAS-SNR and SEP for simulated individual NIR models (blue), simulated individual Raman spectra (red) and simulated concatenated Raman-NIR models (green).

CHAPTER IV

FUTURE WORK

The research in this thesis focuses on combining NIR and Raman spectra in an attempt to improve the prediction ability of resulting PLS calibration models for the measurement of glucose in aqueous solution. The results presented in Chapter II illustrate that PLS calibration models can be improved, but only when the combined spectra provide an improvement in the overall SNR of the analyte specific multivariate calibration vector. For the spectral data presented in Chapter II, the NIR spectral data has a higher NAS-SNR compared to the Raman spectral data. In this case, combining the Raman and NIR data sets results in SEC and SEP values in between those produced from models generated from the individual Raman and NIR data sets alone. The simulated data presented in Chapter III suggest that combining NIR and Raman spectra with the same NAS-SNR can improve the prediction ability of the corresponding PLS model. The relationship between the NAS-SNR and SEP of PLS calibration models is defined in both Chapters II and III for both actual experimental data and simulated data.

The sample matrix for the experiments performed in this thesis consisted of three components (glucose, urea, and lactate) in an aqueous solvent. The generality of the conclusions noted above must be verified in other, more complex matrices, such as whole blood. Because the sample matrix was fairly simple for the work described in Chapters II and III, the SEC and SEP values for the PLS models are low. By exploring more complex matrices, more challenging PLS models will be required, and the benefits of combining spectral data types might be accentuated.

Strategies for combining the two techniques can be expanded further and the benefits of searching for optimized spectral ranges within the combined NIR and Raman spectra might prove useful. In addition, the underlying problem for the real spectra data presented in Chapter II was the NAS-SNR of the Raman spectra is lower than that for the

NIR spectra. Improvements can be made to reduce the noise level of the Raman spectroscopy.

The ultimate goal is to obtain noninvasive measurements of glucose in human subjects. If the combination of Raman and NIR spectra is to be realized for this application, the spectral characteristics of human skin must be established. Recently, the characterization of rat and human skin has been reported for NIR spectroscopy.[13] These same types of skin characterization experiments must be performed with Raman spectroscopy and the resulting NAS-SNR values compared. As indicated in this thesis, the NAS-SNR must be comparable to realize any benefits from combining these spectral types for noninvasive glucose measurements. For subcutaneous Raman measurements, the impact of auto-fluorescence from skin and blood on the NAS-NSR for glucose must be established. Alternatively, ultrafiltration probes can be used to filter out larger molecules in blood, thereby simplifying Raman spectra, improving NAS-SNR, and enabling the use of portable Raman instrumentation.[50]

REFERENCES

1. Alberti, K. G. M. M.; Zimmet, P. Z., Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus. Provisional report of a WHO Consultation. *Diabetic Medicine* **1998,** *15* (7), 539-553.

2. World Health Organization. Diabetes Programme. http://www.who.int/diabetes/en/ (accessed Jan 28, 2013).

3. Amos, A. F.; McCarty, D. J.; Zimmet, P., The Rising Global Burden of Diabetes and its Complications: Estimates and Projections to the Year 2010. *Diabetic Medicine* **1997,** *14* (S5), S7-S85.

4. Khalil, O. S., Spectroscopic and Clinical Aspects of Noninvasive Glucose Measurements. *Clinical Chemsitry* **1999,** *45* (2), 165-177.

5. Newman, J. D.; Turner, A. P. F., Home Blood Glucose Biosensors: A Commercial Perspective. *Biosensors and Bioelectronics* **2005,** *20* (12), 2435-2453.

6. Trief, P. M.; Wade, M. J.; Britton, K. D.; Weinstock, R. S., A Prospective Analysis of Marital Relationship Factors and Quality of Life in Diabetes. *Diabetes Care* **2002,** *24* (7), 1621-1627.

7. Burmeister, J. J.; Arnold, M. A., Evaluation of Measurement Sites for Noninvasive Blood Glucose Sensing with Near-Infrared Transmission Spectroscopy. *Clinical Chemsitry* **1999,** *45* (9), 1621-1627.

8. Enejder, A. M. K.; Scecina, T. G.; Oh, J.; Hunter, M.; Shih, W.-C.; Sasic, S.; Horowitz, G. L.; Feld, M. S., Raman spectroscopy for noninvasive glucose measurements. *Journal of Biomedical Optics* **2005,** *10* (3), 031114-031114.

9. Ferrante do Amaral, C. E.; Wolf, B., Current development in non-invasive glucose monitoring. *Medical Engineering & Physics* **2008,** *30* (5), 541-549.

10. Larin, K. V.; Ghosn, M. G.; Ivers, S. N.; Tellez, A.; Granada, J. F., Quantification of glucose diffusion in arterial tissues by using optical coherence tomography. *Laser Physics Letters* **2007,** *4* (4), 312-317.

11. Weiss, R. Y., Y.; Shusterman, A.; Raz, I., Noninvasive Continuous Glucose Monitoring Using Photoacoustic Technology-Results from the First 62 Subjects. *Diabetes Technology and Therapeutics* **2007,** *9* (1), 68-74.

12. Robinson, M. R. E., R.P.; Haaland, D.M.; Koepp, G.W.; Thomas, E.V.; Stallard, B.R.; Robinson, P.L., Noninvasive Glucose Monitoring in Diabetic Patients: A Preliminary Evaluation. *Clinical Chemistry* **1992,** *38* (9), 1618-1622.

13. Alexeeva, N. V. A., M.A., Near-Infrared Microspectroscopic Analysis of Rat Skin Tissue Heterogeneity in Relation to Noninvasive Glucose Sensing. *Journal of Diabetes Science and Technology* **2009,** *3* (2), 219-232.

14. Chen, J.; Arnold, M. A.; Small, G. W., Comparison of Combination and First Overtone Spectral Regions for Near-Infrared Calibration Models for Glucose and Other Biomolecules in Aqueous Solutions. *Analytical Chemistry* **2004,** *76* (18), 5405-5413.

15. Yonzon, C. R.; Haynes, C. L.; Zhang, X.; Walsh, J. T.; Van Duyne, R. P., A Glucose Biosensor Based on Surface-Enhanced Raman Scattering: Improved Partition Layer, Temporal Stability, Reversibility, and Resistance to Serum Protein Interference. *Analytical Chemistry* **2003,** *76* (1), 78-85.

16. Dou, X.; Yamaguchi, Y.; Yamamoto, H.; Uenoyama, H.; Ozaki, Y., Biological Applications of Anti-Stokes Raman Spectroscopy: Quantitative Analysis of Glucose in Plasma and Serum by a Highly Sensitive Multichannel Raman Spectrometer. *Applied Spectroscopy* **1996,** *50* (10), 1301-1306.

17. Berger, A. J.; Koo, T.-W.; Itzkan, I.; Horowitz, G.; Feld, M. S., Multicomponent Blood Analysis by Near-Infrared Raman Spectroscopy. *Appl. Opt.* **1999,** *38* (13), 2916-2926.

18. Berger, A. J.; Itzkan, I.; Feld, M. S., Feasibility of measuring blood glucose concentration by near-infrared Raman spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **1997,** *53* (2), 287-292.

19. Enejder, A. M. K.; Koo, T.-W.; Oh, J.; Hunter, M.; Sasic, S.; Feld, M. S.; Horowitz, G. L., Blood analysis by Raman spectroscopy. *Opt. Lett.* **2002,** *27* (22), 2004-2006.

20. Burns, D. A. C., E.W., *Handbook of Near-Infrared Analysis, 2nd edition*. Marcel Dekker, Inc. : 2001.

21. Ren, M. Comparison of near infrared and Raman spectroscopies for noninvasive clinical measurements. The University of Iowa, 2007.

22. Albergotti, J. C., Fourier Transform Spectroscopy Using a Michelson Interferometer. *American Journal of Physics* **1972,** *40* (8), 1070-1078.

23. Cochran, W. T.; Cooley, J. W.; Favin, D. L.; Helms, H. D.; Kaenel, R. A.; Lang, W. W.; Maling, G. C., Jr.; Nelson, D. E.; Rader, C. M.; Welch, P. D., What is the fast Fourier transform? *Proceedings of the IEEE* **1967,** *55* (10), 1664-1674.

24. Perkins, W. D., Fourier transform infrared spectroscopy. Part II. Advantages of FT-IR. *Journal of Chemical Education* **1987,** *64* (11), A269.

25. Skoog, D. A.; Holler, J. M.; Nieman, T. A., *Principles of Instrumental Analysis, Fifth Edition*. Thomson Learning, Inc.: 1998.

26. Pelletier, M. J., Quantitative Analysis Using Raman Spectrometry. *Applied Spectroscopy* **2003,** *57* (1), 20A-42A.

27. Schulz, H.; Baranska, M., Identification and quantification of valuable plant substances by IR and Raman spectroscopy. *Vibrational Spectroscopy* **2007,** *43* (1), 13-25.

28. Bell, S. E. J.; Bourguignon, E. S. O.; Dennis, A., Analysis of luminescent samples using subtracted shifted Raman spectroscopy. *Analyst* **1998,** *123* (8), 1729-1734.

29. O'Grady, A.; Dennis, A. C.; Denvir, D.; McGarvey, J. J.; Bell, S. E. J., Quantitative Raman Spectroscopy of Highly Fluorescent Samples Using Pseudosecond Derivatives and Multivariate Analysis. *Analytical Chemistry* **2001,** *73* (9), 2058-2065.

30. Lavine, B.; Workman, J. J., Chemometrics. *Analytical Chemistry* **2004,** *76* (12), 3365-3372.

31. Bro, R., Multivariate calibration: What is in chemometrics for the analytical chemist? *Analytica Chimica Acta* **2003,** *500* (1–2), 185-194.

32. Sijmen, d. J., SIMPLS: An alternative approach to partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* **1993,** *18* (3), 251-263.

33. Haaland, D. M.; Thomas, E. V., Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry* **1988,** *60* (11), 1193-1202.

34. Haaland, D. M.; Thomas, E. V., Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data. *Analytical Chemistry* **1988,** *60* (11), 1202-1208.

35. Amerov, A. K.; Chen, J.; Arnold, M. A., Molar Absorptivities of Glucose and Other Biological Molecules in Aqueous Solutions over the First Overtone and Combination

Regions of the Near-Infrared Spectrum. *Applied Spectroscopy* **2004,** *58* (10), 1195-1204.

36. Amerov, A. K.; Chen, J.; Small, G. W.; Arnold, M. A., Scattering and Absorption Effects in the Determination of Glucose in Whole Blood by Near-Infrared Spectroscopy. *Analytical Chemistry* **2005,** *77* (14), 4587-4594.

37. Lorber, A., Error propagation and figures of merit for quantification by solving matrix equations. *Analytical Chemistry* **1986,** *58* (6), 1167-1172.

38. Lorber, A.; Faber, K.; Kowalski, B. R., Net Analyte Signal Calculation in Multivariate Calibration. *Analytical Chemistry* **1997,** *69* (8), 1620-1626.

39. Ren, M.; Arnold, M., Comparison of multivariate calibration models for glucose, urea, and lactate from near-infrared and Raman spectra. *Anal Bioanal Chem* **2007,** *387* (3), 879-888.

40. Arnold, M. A.; Small, G. W., Noninvasive Glucose Sensing. *Analytical Chemistry* **2005,** *77* (17), 5429-5439.

41. Cai, W.; Li, Y.; Shao, X., A variable selection method based on uninformative variable elimination for multivariate calibration of near-infrared spectra. *Chemometrics and Intelligent Laboratory Systems* **2008,** *90* (2), 188-194.

42. Lin-Vien, D., *The Handbook of infrared and raman characteristic frequencies of organic molecules*. Boston : Academic Press: 1991.

43. Amerov, A. K.; Arnold, M. A., In Vitro Kromoscopic Analysis of Glucose in Blood. *Proceedings of SPIE* **2003,** *4965*, 7-16.

44. Mathlouthi, M.; Vinh Luu, D., Laser-raman spectra of d-glucose and sucrose in aqueous solution. *Carbohydrate Research* **1980,** *81* (2), 203-212.

45. Lambert, J. L.; Morookian, J. M.; Sirk, S. J.; Borchert, M. S., Measurement of aqueous glucose in a model anterior chamber using Raman spectroscopy. *Journal of Raman Spectroscopy* **2002,** *33* (7), 524-529.

46. Hoccart, X.; Turrell, G., Raman spectroscopic investigation of the dynamics of urea-water complexes. *Journal of chemical Physics* **1993,** *99* (11), 8498-8503.

47. Frost, R. L.; Kristof, J.; Rintoul, L.; Kloprogge, J. T., Raman spectroscopy of urea and urea-intercalated kaolinites at 77 K. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy* **2000,** *56* (9), 1681-1691.

48. Pecul, M.; Rizzo, A.; Leszczynski, J., Vibrational Raman and Raman Optical Activity Spectra of d-Lactic Acid, d-Lactate, and d-Glyceraldehyde: Ab Initio Calculations. *The Journal of Physical Chemistry A* **2002,** *106* (46), 11008-11016.

49. Cassanas, G.; Kister, G.; Fabrègue, E.; Morssli, M.; Bardet, L., Raman spectra of glycolic acid, l-lactic acid and d,l-lactic acid oligomers. *Spectrochimica Acta Part A: Molecular Spectroscopy* **1993,** *49* (2), 271-279.

50. Hargreaves, M. D.; Page, K.; Munshi, T.; Tomsett, R.; Lynch, G.; Edwards, H. G. M., Analysis of seized drugs using portable Raman spectroscopy in an airport environment—a proof of principle study. *Journal of Raman Spectroscopy* **2008,** *39* (7), 873-880.