Comparison of mid-infrared and Raman spectroscopy in the quantitative analysis of serum

D. Rohleder
Roche Diagnostics GmbH
Mannheim, Germany
and
University of Würzburg
Institute for Physical Chemistry
Würzburg, Germany

G. Kocherscheidt
K. Gerber
Dioptic GmbH
Weinheim, Germany

W. Kiefer
University of Würzburg
Institute for Physical Chemistry
Würzburg, Germany

W. Köhler
BaseLine GbR
Frankfurt, Germany

J. Möcks
Roche Diagnostics GmbH
Mannheim, Germany

W. Petrich
Roche Diagnostics GmbH
Mannheim, Germany
and
University of Heidelberg
Faculty of Physics and Astronomy
Heidelberg, Germany

Abstract. Mid-infrared or Raman spectroscopy together with multivariate data analysis provides a novel approach to clinical laboratory analysis, offering benefits due to its reagent-free nature, the speed of the analysis and the possibility of obtaining a variety of information from one single measurement. We compared mid-infrared and Raman spectra of the sera obtained from 247 blood donors. Partial least squares analysis of the vibrational spectra allowed for the quantification of total protein, cholesterol, high and low density lipoproteins, triglycerides, glucose, urea and uric acid. Glucose (mean concentration: 154 mg/dl) is frequently used as a benchmark for spectroscopic analysis and we achieved a root mean square error of prediction of 14.7 and 17.1 mg/dl for mid-infrared and Raman spectroscopy, respectively. Using the same sample set, comparable sample throughput, and identical mathematical quantification procedures Raman and mid-infrared spectroscopy of serum deliver similar accuracies for the quantification of the analytes under investigation. In our experiments vibrational spectroscopy-based quantification appears to be limited to accuracies in the 0.1 mmol/l range. [DOI: 10.1117/1.1911847]

Keywords: Mid-infrared spectroscopy; Raman spectroscopy; serum.

Paper SS04144RR received Jul. 28, 2004; revised manuscript received Dec. 23, 2004; accepted for publication Dec. 28, 2004; published online May 24, 2005.

1 Introduction

The spectroscopy of molecular vibrations is experiencing a renaissance due to substantial technical advances in experimental methods, increased computational capabilities and growing analytical demands. Potential medical applications are currently being investigated in various fields such as angiology, rheumatology, endocrinology, dentistry, or dermatology.[1–5] In a routine clinical laboratory setting, body fluids like blood are particularly easy to obtain and are thus the most frequent primary material of investigation by laboratory diagnostics.

Blood has frequently been subjected to infrared and Raman spectroscopy.[6–9] Since cellular components such as erythrocytes account for ~42% of the weight of blood, the vibrational spectroscopy of whole blood is modified by Mie scattering. Removal of the solid components leaves the liquid phase, which is called plasma. Although many investigations have shown very promising results for plasma,[10–13] the collection of plasma unfortunately requires the addition of a highly standardized type and concentration of anticoagulant. In contrast, serum is readily obtained from whole blood if (in addition to the cellular components) those substances are removed which contribute to the coagulation cascade, e.g., fibrinogen. In daily routine, serum is simply collected from whole blood by means of centrifugation. Hence, this manuscript focuses on the serum which is an easily accessible sample, needs no addition of anticoagulant, can be obtained reproducibly and is a very frequent sample type in routine laboratory diagnostics.

Water forms the basis of body fluids and accounts for 90% of the serum. In the mid-infrared region, where most of the basic vibrations of biomolecules occur, the absorption coefficient of water amounts to 10^2-10^4/cm making the direct observation of vibrational modes of organic compounds in body fluids difficult. This difficulty may be mitigated or bypassed, e.g., by using transmission path lengths below 100 μm[14] and/or intense mid-infrared light sources,[15,16] by employing attenuated total reflection techniques,[10] or by investigating near-infrared overtones instead of the fundamental modes of the vibrations.[11,17] Moreover, the body fluid could be dried such that the sample's water content is substantially reduced and the remaining components of the body fluid can be di-

Address all correspondence to Dr. Wolfgang Petrich, Roche Diagnostics, Sanhofer Str. 116, Mannheim, 68305 Germany. Tel: +49-621-759-8324. E-mail: wolfgang.petrich@roche.com

1083-3668/2005/$22.00 © 2005 SPIE

Table 1 Mean, maximum and minimum values and standard deviation σ_{ref} of the different analyte concentrations as determined by the reference method. The relative coefficient of variation (%CV) measures the precision of the reference methods upon re-sampling. The concentration of LDL cholesterol was determined using the Friedewald formula.

Analyte	Mean [mg/dl]	Min [mg/dl]	Max [mg/dl]	σ_{ref} [mg/dl]	Test method	%CV
Total protein	7008	6100	8100	376	Colorimetric assay	0.95
Glucose	154	42	423	103	Enzymatic UV test	1.7
Urea	31	15	56	7	Kinetic UV assay	3.4
Uric acid	5.3	2.5	9.0	1.3	Enzymatic colorimetric test	1.7
Cholesterol	133	37	392	38	CHOD-PAP	1.7
Triglycerides	198	119	338	68	GPO-PAP	1.8
HDL cholesterol	54	14	99	14	Enzymatic colorimetric test	1.85
LDL cholesterol	118	47	249	34	…	…

rectly investigated by mid-infrared spectroscopy.[6] Alternatively, various kinds of Raman spectroscopy may be used to provide access to the fundamental vibrational spectra of body fluids.[8,9,18–20] Near-infrared Raman spectroscopy appears to be particularly favorable, since it takes advantage of the low absorption coefficient of water in the near-infrared spectral region (which, e.g., amounts to 0.35/cm at a wavelength of 1 μm) and since the fluorescent light background is strongly reduced when compared to using visible light for Raman spectroscopy.

For the research described in this report, two approaches were followed in order to investigate body fluids, namely near-infrared Raman spectroscopy of serum in its native form and mid-infrared spectroscopy of dried films of serum. While each of the two spectroscopic techniques was optimized individually, identical samples were used for both analyses, the samples throughput was required to be identical and the final data analysis was performed with the identical software routines in order to allow for a close comparison of the two approaches.

2 Materials and Methods

Blood samples were collected from 238 healthy donors and nine patients suffering from diabetes. The samples were centrifuged at 900 g for 30 min using a Heraeus Labofuge GL and the serum was isolated. Serum samples from 80 donors were set aside and glucose was added to these samples in the form of a glucose solution (Fresenius Kabi Glucosteril® 70%). Forty of these samples were spiked with 2.2 μL glucose solution per millilitre of serum to increase the glucose concentrations by approximately 150 mg/dl. Glucose concentrations in the other 40 samples were increased by approximately 300 mg/dl upon addition of 4.4 μL glucose solution per millilitre of serum. Subsequently, all of the serum samples were partitioned into multiple aliquots of 3 mL each. All samples were frozen at −80 °C for storage purposes. One of the aliquots of each donor's samples was subjected to clinical chemistry testing: The concentrations of total protein, glucose, uric acid, urea, cholesterol, triglycerides and HDL cholesterol were determined by enzymatic tests using a MODULAR® PP system.[§3] The concentration of LDL cholesterol was calculated by the Friedewald formula. Note that these reference concentrations of analytes were determined after the aforementioned steps, particularly after spiking. Furthermore, the concentrations of cholesterol, triglycerides, HDL and LDL are physiologically interrelated. While the square of Pearson's correlation coefficient indicates a substantial correlation ($r^2 = 0.86$) between cholesterol and LDL, r^2 is less than 0.3 for all other pairs of metabolites. For example, triglycerides are not completely unrelated to cholesterol ($r^2 = 0.19$) and HDL ($r^2 = 0.21$). Qualitatively, these observations also hold true when calculating r^2 within the various subsets (e.g., spiked samples versus unspiked samples) individually.

Minimum, mean, and maximum concentrations of each analyte are given in Table 1 together with the standard deviation of the concentrations and the reference test method. The standard measure for precision in a clinical laboratory is the coefficient of variation, which is determined by remeasuring the concentrations of analytes multiple times. The ratio between the coefficient of variation and the mean value of the concentration of the analyte under investigation, i.e., the relative coefficient of variation %CV, is listed in Table 1 for each of the individual reference methods.

Since more than 180 million people suffer from diabetes mellitus, a disorder of the glucose metabolism, the quantification of glucose has frequently served as a benchmark for the capabilities of vibrational spectroscopy in the context of clinical laboratory analysis. The quantification of glucose has therefore been emphasized throughout this manuscript.

We employed an automated pipetting system together with a BRUKER Matrix HTS-XT spectrometer for the experiments using mid-infrared spectroscopy.[21] Ninety-six-well silicon sample carriers were used for mid-infrared transmission spectroscopy; 3 μL of serum were pipetted onto a sample carrier in random order and left to dry in ambient air for 30 min. After drying, the film (thickness of 2–10 μm) was subjected

[§]MODULAR is a trademark of a member of the Roche group.

to mid-infrared transmission spectroscopy. Spectroscopy is performed in transmission using a DLaTGS detector, which in contrast to mercury cadmium telluride (MCT) detectors can be operated without liquid nitrogen cooling. Each spectrum was recorded in the wave number range from 500 to 4000 cm^{-1} and consisted of 3629 data points. Spectra were acquired at a resolution of 4 cm^{-1} and averaged over 32 scans. Blackmann–Harris three-term apodization was used and the zero filling factor was 4 (note that the use of a zero filling of 4 is standard practice in our laboratory but does not contribute to the accuracy of our results). To improve reproducibility each sample was pipetted and measured on three sample carriers and the three absorbance spectra of each sample were corrected for sample carrier background and normalized. We used a proprietary algorithm[22] to correct for sample carrier background and to normalize the spectra. (Note that due to the good reproducibility of our measurement method during this study a simple background subtraction and standard vector normalization will give similar results). At each wave number the median of the three pre-processed spectra was calculated and the resulting spectrum was then subjected to further analysis. Although the actual integration time per spectrum was only 29 s, the triplicate measurement, the spectroscopic determination of the sample carrier background and the sample handling time resulted in an average processing time of 5 min per sample. The sample carriers were discarded after use.

Stability of the mid-infrared system was coarsely assessed by calculating the area under the spectra before normalizing. Variations of the area under each spectrum are caused primarily by variations of the shape and thickness of the dried film of serum, which finally leads to a variation in optical path length. We find that the area under curve varied by less than $\pm 20\%$ among the measurements in this study. However, one sample was excluded from further analysis since the area under the mid-infrared spectra amounted to less than 50% of the expected value for all three absorbance spectra originating from this sample.

A Kaiser Optical HoloSpec f/1.8i spectrometer was used for Raman spectroscopy. Laser radiation (wavelength 785 nm) interacted with the sample within a quartz cuvette (power at the location of the sample: 200 mW) and backscattered radiation was collected using an Olympus PL4X lens. Ten quartz cuvettes were alternated and, after ten measurements, the cuvettes were cleaned in 1% Hellmanex II solution (Hellma GmbH&Co. KG, Müllheim/Baden, Germany) at 70 °C, dried and used for the next set of measurements. Spectral resolution was 8 cm^{-1}. Spectra were acquired over 5 min during 12 acquisitions of 25 s each. The raw spectra were normalized and a fifth order polynomial background was subtracted in the region from 300 to 1870 cm^{-1} using an iterative algorithm.[23] Further details of the Raman experiments are reported in Ref. 19.

The strategy of the comparison was to use an optimum setting for each spectroscopic method individually, but to require the working conditions from a laboratory standpoint and the data analysis to be as equivalent as possible for both methods. Differences and similarities between the parameters used for the two approaches are listed in Table 2. While many of the parameters had become an internal working standard during our prior investigations, we paid particular attention to request a throughput of at least 80 samples per day for both spectroscopic methods. Furthermore, liquid nitrogen cooling had to be avoided in the view of possible future laboratory application. We used samples from the same study for the investigation of both spectroscopic approaches and we split the spectra into the same calibration and validation sample subsets. Finally, identical multivariate analysis algorithms were used for the quantitative analysis of the pre-processed spectra.

After the spectroscopy and the pre-processing of spectra had been completed, all data (laboratory data and spectra) of the 247 donors were divided into a teaching set of 148 donors' data and a set of 99 donors' data for independent validation. Those samples exhibiting the lowest and the highest concentrations of the different analytes were always assigned to the teaching set. The statistical equivalence of the teaching and validation sets was verified on the basis of two-sample *t*-tests and two-sample *F*-tests for the different analytes. For teaching, partial least squares regression (PLS) was performed using MathWorks' MatLab™ 6.0 Release 12 together with SIMPLS algorithm implemented in the PLS_Toolbox 2.1 by Eigenvector Research, Inc. In order to optimize the training within the teaching data set, the root-mean-square error of calibration (RMSEC) and the root-mean-square error of leave-one-out cross validation (RMSECV) were calculated

$$\mathrm{RMSEC}=\left[\frac{\sum_{i=1}^{N_{\mathrm{teach}}}(c_{\mathrm{pred},i}-c_{\mathrm{ref},i})^2}{N_{\mathrm{teach}}-LV-1}\right]^{1/2}$$

$$\mathrm{RMSECV}=\left[\frac{\sum_{i=1}^{N_{\mathrm{teach}}}(c_{\mathrm{pred},i}-c_{\mathrm{ref},i})^2}{N_{\mathrm{teach}}}\right]^{1/2}.$$

Here $c_{\mathrm{ref},i}$ and $c_{\mathrm{pred},i}$ denote the concentrations of analytes in sample i as determined by the reference method and by the spectroscopic measurement, respectively. N_{teach} is the number of teaching samples ($N_{\mathrm{teach}}=148$) and LV is the number of latent variables used for the PLS calibration. The optimum LV was chosen by selecting that value for LV, which corresponds to the minimum of RMSECV.

The validation set remained blinded until the teaching had been finalized. As a measure for the prediction accuracy of the system, the root-mean-square error of prediction (RMSEP) was calculated according to

$$\mathrm{RMSEP}=\left[\frac{\sum_{i=1}^{N_{\mathrm{val}}}(c_{\mathrm{pred},i}-c_{\mathrm{ref},i})^2}{N_{\mathrm{val}}}\right]^{1/2}.$$

N_{val} is the number of validation samples ($N_{\mathrm{val}}=99$ for Raman spectroscopy, $N_{\mathrm{val}}=98$ for mid-infrared spectroscopy). Relative errors (%RMSEP) are calculated as the ratio between RMSEP and the mean concentration.

3 Results

An example of the mid-infrared spectrum of a dried film of serum is given in Fig. 1. The mid-infrared spectrum is dominated by the infrared absorption of proteins such as albumin or globulins, which, after drying, constitute the major components of the serum film. Proteins exhibit characteristic vibrations of the polypeptide skeleton. The most pronounced peak

Table 2 Main characteristics of the parameters used for infrared and Raman spectroscopy. Note that aliquots of identical serum samples have been used for both approaches.

	FTIR	RAMAN
Process parameters		
Samples throughput	80/day	
Need for liquid nitrogen cooling	no	
Sample carrier type	Silicon plate	Quartz cuvette
Sample carrier reuse	No	Yes
Background measurement	Yes	No
Sample volume used	100 μL	1 mL
Sample handling	Automated	Manual
Sample drying	Yes	No
Multiplicity of measurement	Triplicate	Single
Spectroscopy parameters		
Light source	Globar	Semiconductor laser
Detector type	DLaTGS	CCD
Acquisition time for a single spectrum	30 s	5 min.
Detected wave number range	500–4000 cm^{-1}	300–3500 cm^{-1}
Spectral resolution	4 cm^{-1}	8 cm^{-1}
Zero filling	4	~2
Analysis parameters		
Data pre-treatment	Background correction, normalization	Subtraction of 5th order polynomial, normalization
Wave number range used for PLS analysis of proteins	1220–1690 cm^{-1}	300–1500 cm^{-1}
Wave number range used for PLS analysis of all other analytes	500–1800 and 2500–3300 cm^{-1}	300–1500 cm^{-1}
Teaching set	148 serum samples	
Teaching algorithm	SIMPLS	
Determination of optimum No. LV	minimum of RMSECV	
Independent validation set	99 serum samples	
Measure of quality of quantification	RMSEP	

at 1653 cm^{-1} is caused by the Amide I vibration of the peptide chain. Similarly, the peaks at 1545 cm^{-1} and around 1270 cm^{-1} can be assigned to the Amide II and Amide III vibrations. The O–H stretch vibration is reflected as a broad feature around 3300 cm^{-1}, which also contains a triplet structure arising from stretch vibrations of $-C-H$ in $-CH_2$ and $-CH_3$.

Figure 2 shows the Stokes-shifted Raman signal of serum after background subtraction. While the $-C-H_x$ stretch vibrations around 2900 cm^{-1} appear to be similar to the mid-infrared case, the spectrum substantially differs at lower wave numbers. The Amide I band is strongly decreased and the Amide III band is part of the most prominent feature of the spectrum. Furthermore, the essential amino acids phenylalanine and tyrosine can be clearly identified at 1003 cm^{-1} and at the 829/851 cm^{-1} duplet, respectively.

Quantification of the concentration of the analytes was performed by training a PLS algorithm individually for each analyte within the spectral regions shown in Table 2. As an example, the root-mean-square errors resulting from the PLS analysis of infrared spectra are illustrated in Fig. 3 for the case

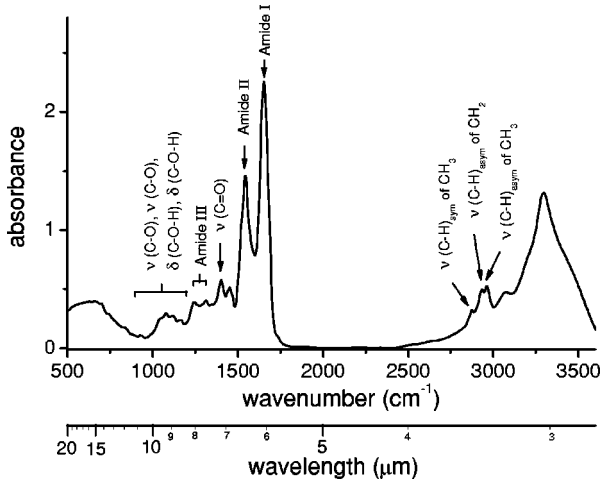

Fig. 1 Mid-infrared spectrum of a film of dried serum after subtraction of the background signal caused by the silicon sample carrier (ν: stretch vibration. δ: bending vibration).

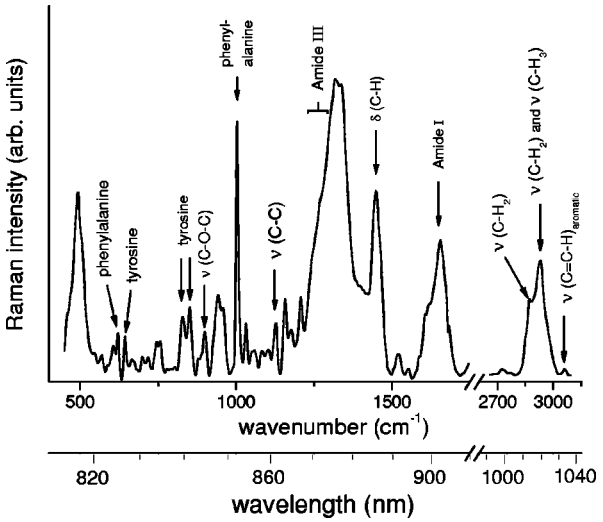

Fig. 2 Raman signal of serum after background subtraction. The wavelength scale represents the wavelength of the Raman scattered light during illumination of the sample with laser light at 785 nm (ν: stretch vibration. δ: bending vibration). The energy difference allows for the calculation of the Raman shift, which is expressed in terms of wave numbers in this graph.

of glucose. Here, the minimum of RMSECV occurs at $LV_{\min}=15$. For glucose as well as the other seven parameters $LV_{\min}$ is given in Table 3 together with the corresponding values for RMSEC and RMSECV.

Since the values of RMSECV are very similar in the vicinity of $LV_{\min}$, it is interesting to ask in how far RMSECV becomes minimal at the given values of $LV_{\min}$ at random. In other words, will the minimum of RMSECV be found at exactly the same values of $LV_{\min}$ as listed in Table 3 if one were to repeat the whole experiment? We find that, for instance, in the case of glucose the distribution of residue of the leave-one-out cross validation for $LV_{\min}=15$ is not significantly different from that calculated for any value of LV between 9 and 31 (F-test, $\alpha=0.05$). Similarly, different ranges of statistically equivalent values for LV are given in Table 3 for the eight parameters under investigation.

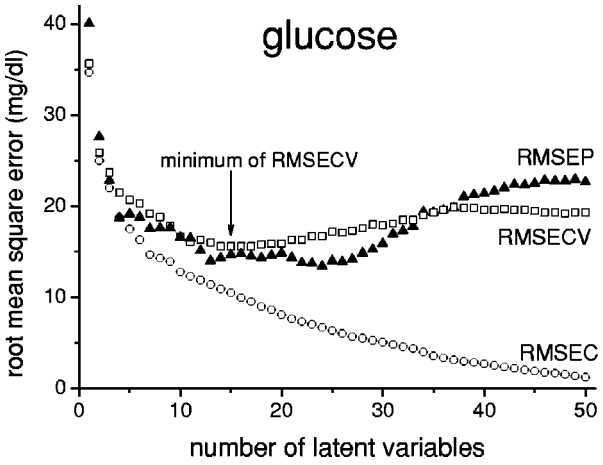

Fig. 3 Root-mean-square error of calibration (RMSEC), leave-one-out cross validation (RMSECV) and prediction (RMSEP) as a function of the number of latent variables using the quantification of glucose based on infrared spectra as an example. The minimum of RMSECV determines the "optimum" number of latent variables (15 in the case of glucose).

After the teaching of the algorithms had been finalized and $LV_{\min}$ had been determined, the spectra of the validation set were subjected to blinded validation. The analyte concentrations of the 98 validation samples investigated by mid-infrared spectroscopy were predicted based on the PLS model and subsequently compared to the concentration derived by the laboratory method (see Fig. 4; the predictions for the case of the 99 Raman spectra are illustrated in Ref. 19). The results of the quantitative analysis are summarized in Table 4 for both, mid-infrared and Raman spectroscopy. For completeness and since the particular values of $LV_{\min}$ are subject to some randomness (as outlined above) we have also calculated the RMSEP for all those PLS models, in which the number of latent variables was within the discussed ranges of statistically equivalent values of LV. The minimum ($\mathrm{RMSEP}_{\min}$) and maximum ($\mathrm{RMSEP}_{\max}$) prediction error observed within the given range of LV are listed in Table 4 for each analyte.

The predictions obtained by using mid-infrared spectroscopy may be compared to the results of Raman spectroscopy. As an example, the difference between the predicted concentration and its actual concentration as determined by laboratory analysis is illustrated for the case of glucose in Fig. 5 for both technologies. Beyond the qualitative impression of the scatter of the data, the calculation of RMSEP allows for a more quantitative comparison. A RMSEP of 14.7 and 17.1 mg/dl was achieved for mid-infrared and Raman spectroscopy, respectively. No significant differences between the mean values of the shown residue (paired t-test; $P=0.03$) as well as their spread (F-test; $P=0.04$) could be detected between the two spectroscopic methods.

4 Discussion

Our comparative study was dedicated to the quantitative analysis of Raman spectra of native serum and mid-infrared spectra of films formed from serum upon drying. To the best of our knowledge the investigations presented in this manuscript constitute the most comprehensive comparison between mid-infrared and Raman spectroscopy with regard to determining the concentration of analytes in serum.

Particular attention was paid to requesting identical operating parameters for both methods from a clinical laboratory viewpoint, namely equal throughput and the avoidance of liquid nitrogen cooling. In addition, the data analysis procedures were identical once the pre-treatment of the raw spectra of each method was finalized. Furthermore, it was important that the identical samples were used for both methods, including the identical splitting and sorting into teaching and validation data sets.

For an unbiased analysis it was important to rigorously train the PLS algorithm using the teaching data only and to perform an independent validation thereafter. As part of this clear separation between the teaching and the validation data set, the search for the optimum dimensionality of the problem (i.e., identifying "optimum" number of latent variables $LV_{\min}$ for the PLS algorithm) was based on the teaching set only. After all parameters of the PLS model had been defined, in-

Table 3 Results of the teaching procedure. LV_{min} denotes that number of latent variables, for which the root-mean-square error of leave-one-out cross validation (RMSECV) becomes minimal. RMSEC is the root-mean-square error of calibration. Since RMSECV does not statistically differ from its minimum value within a range of latent variables, the range of statistically equivalent values of LV is noted in brackets.

	LV_{min}		RMSEC [mg/dl]		RMSECV [mg/dl]	
Analyte	IR	Raman	IR	Raman	IR	Raman
Total protein	3 (1–20)	10 (7–15)	300	118	318	157
Glucose	15 (9–31)	10 (8–15)	9.5	17.1	15.6	22.6
Urea	18 (11–50)	12 (11–19)	2.3	2.5	4.0	3.9
Uric acid	10 (1–19)	12 (1–28)	0.9	0.6	1.2	1.2
Cholesterol	12 (8–25)	12 (10–50)	10.9	7.4	15.1	11.8
Triglycerides	19 (12–28)	15 (10–50)	9.8	7.3	16.7	19.3
HDL cholesterol	15 (4–50)	10 (7–26)	6.6	7.2	10.9	9.8
LDL cholesterol	12 (6–33)	14 (10–50)	13.9	5.9	19.5	15.3

dependent validation was performed and the root-mean-square error of prediction (RMSEP) was calculated. In retrospect we find that those values of LV which provide the minimum value for RMSEP within the independent validation set (see Table 4) are frequently very close to our estimates LV_{min}, which were derived from the teaching set only. Thus, we conclude that the method we used for estimating the optimum number of latent variables provides a reasonable approach to the problem of dimensionality.

In our measurement setup, Raman spectroscopy required larger sample volumes than the infrared spectroscopy. Although we envisage that the volume used here (1 mL) can be reduced to 200 μL by means of automation, it would still be twice the volume used in infrared spectroscopy. For infrared spectroscopy, the volume may even be reduced further: in fact, we designed our system such that it can operate with volumes as low as 70 μL and even lower volumes are conceivable.

Given the fact that vibrational changes in dipole moment (or polarizability in the case of Raman spectroscopy) are of a similar order of magnitude for most biomolecules, the sample-specific detection capabilities mainly depend on the concentration. The RMSEP values of the eight analytes under investigation are shown for both spectroscopic techniques as a function of mean concentration in Fig. 6. RMSEP appears to increase with analyte concentration. However, the ratio between RMSEP and mean value decreases with increasing concentration (dashed lines in Fig. 6): Uric acid exhibits the lowest concentration of all of the analytes investigated and pertains a relative error of up to 26% upon quantification. In contrast, proteins constitute the molecular group with the highest concentration and they can be quantified within a relative error as low as 2.5%. This tendency holds true for both mid-infrared and Raman spectroscopy. In order to relate our findings with present day clinical chemical analyzers it is also instructive to understand the measurement accuracy in terms of the number of molecules rather than their mass-related concentration: Considering the molar weights of the analytes investigated, vibrational-spectroscopy based quantification appears to be limited to accuracies in the 0.1 mmol/L range, regardless of the particular choice of the spectroscopic technique. This finding is also supported by prior publications of our and other groups as listed in Table 5.[6,8,18–19,24–27]

High signal-to-noise ratios are considered a fundamental strength of infrared spectroscopy when compared to Raman spectroscopy. However, we find that this advantage does not result in a superior prediction accuracy when compared to Raman spectroscopy. This result supports our prior finding, that reproducibility rather than signal-to-noise ratio imposes a lower limit on the prediction errors in mid-infrared spectroscopy, even if particular attention is paid to the reproducibility by virtue of automation, triplicate measurement, standardization, and computational efforts.[21] A small supplementary investigation also points at the importance of reproducibility: five randomly chosen samples from the above study were remeasured over the course of the above experiments using mid-infrared spectroscopy and the concentrations of analytes were predicted on the basis of the PLS algorithm described above. For each sample and each analyte, the predicted concentrations vary from measurement to measurement. In analogy to the clinical laboratory guidelines, the relative coefficient of variation (%CV) can be calculated as a measure for

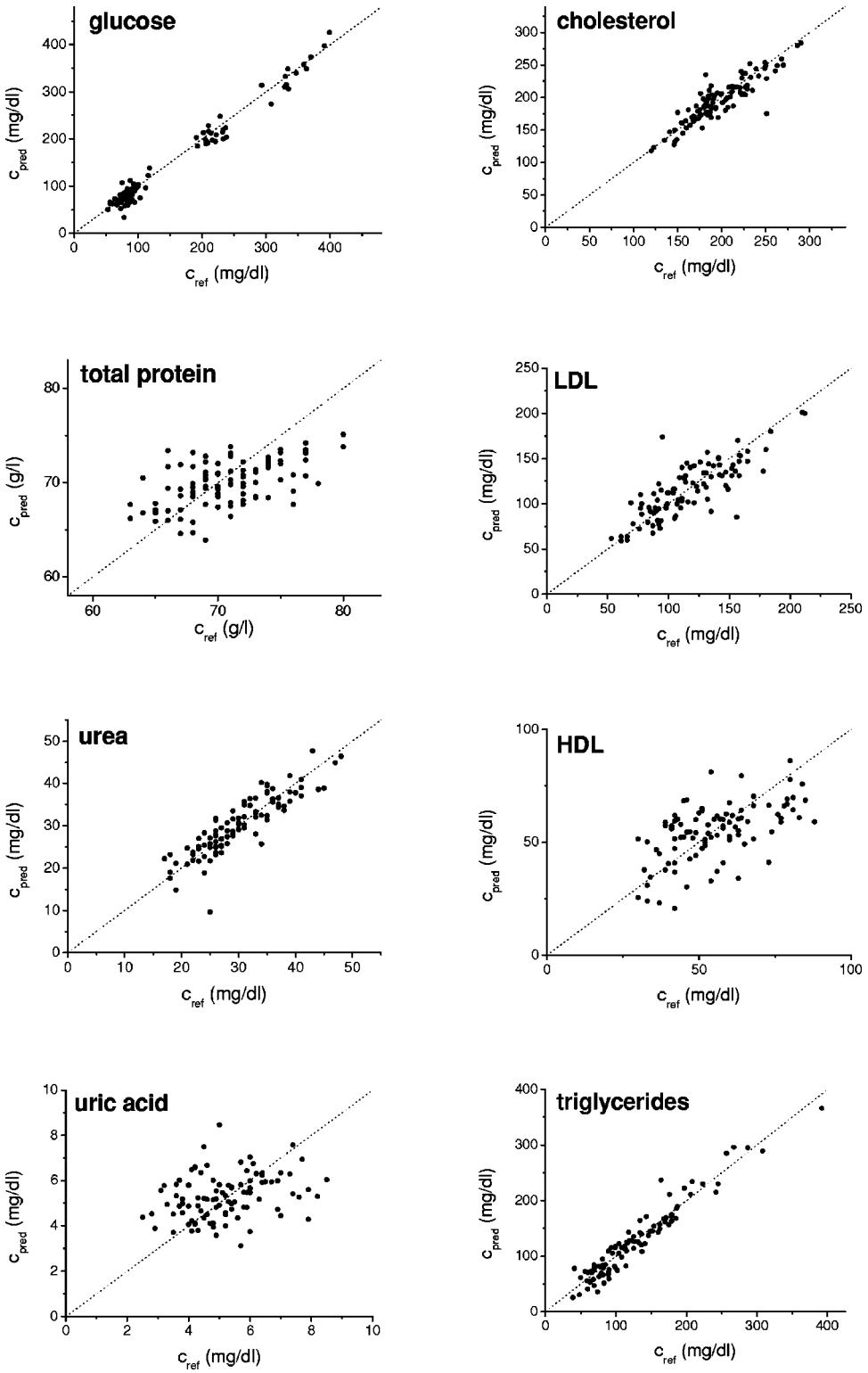

Fig. 4 Concentrations of analytes in the validation samples as predicted by mid-infrared spectroscopy (c_{pred}) as compared to the concentrations determined by the laboratory methods (c_{ref}). Corresponding Raman data have been published in Ref. 19.

the precision of the system. We find that, on average, %CV ranges from 4% (protein) to 16% (LDL). These numbers have to be compared to 4.7% and 16.4% for the %RMSEP of protein and LDL, respectively. Thus, we find that for these, as well as most of the other analytes, the error observed upon remeasuring the sample still substantially contributes to the overall error in the case of infrared spectroscopy of serum. The challenge in reproducibility might be caused by the high susceptibility of mid-infrared spectroscopy to changes in environmental conditions (in particular water vapor and tem-

Table 4 Results of the independent validation. RMSEP{LV_{min}} denotes the root-mean-square error of prediction at that number of latent variables, for which the root-mean-square error of leave-one-out cross validation (RMSECV) became minimal (see Table 3). $RMSEP_{min}$ and $RMSEP_{max}$ denote the minimum and maximum values of RMSEP observed when using PLS calibration models on the basis of different values of LV (in braces) within a range statistically equivalent to LV_{min}.

	RMSEP{LV_{min}} [mg/dl]		$RMSEP_{min}${LV} [mg/dl]		$RMSEP_{max}${LV} [mg/dl]	
Analyte	IR	Raman	IR	Raman	IR	Raman
Total protein	328 {3}	176 {10}	323 {4}	169 {8}	434 {20}	198 {7}
Glucose	14.7 {15}	17.1 {10}	13.4 {24}	16.9 {9}	17.6 {9}	21.1 {14}
Urea	3.3 {18}	4.4 {12}	3.3 {21}	4.4 {12}	5.6 {50}	4.9 {17}
Uric acid	1.4 {10}	1.1 {12}	1.3 {7}	1.1 {11}	1.6 {19}	1.3 {1}
Cholesterol	16.1 {12}	11.5 {12}	15.0 {11}	11.1 {11}	18.0 {24}	14.1 {29}
Triglycerides	18.1 {19}	20.7 {15}	17.5 {17}	19.8 {12}	21.4 {27}	23.9 {50}
HDL cholesterol	11.9 {15}	11.0 {10}	11.8 {14}	10.0 {12}	21.1 {44}	13.7 {25}
LDL cholesterol	19.4 {12}	15.7 {14}	18.6 {18}	14.6 {11}	25.3 {33}	19.1 {50}

perature) which affect both the spectroscopy and the drying process. In turn, the lower signal-to-noise ratio generally observed during Raman spectroscopy does not prevent the quantification of analytes in serum if a measurement time of 5 min per sample is acceptable.

In the light of a routine clinical laboratory application, the relative prediction errors (%RMSEP) may be compared to the standard deviations of reference concentrations, which primarily reflect the physiological variations within the population under investigation. For example, the standard deviation of the reference values amounts to only 5.4% of their mean concentration for total protein. On the other hand, the concentration of proteins can be predicted with a relative prediction error of 4.7% for mid-infrared spectroscopy. Ignoring any non-Gaussian contribution to the distribution of concentrations, it appears reasonable to conclude that the relative prediction error of the infrared spectroscopic approach is comparable to the biological variations of the concentration of total protein in our study population. Similar conclusions may be drawn for HDL and uric acid, for which the relative prediction errors exceed the biological variation among the donors of our study population by only 30% or less. In contrast, the %RMSEP values for cholesterol, triglycerides, LDL and urea are up to four times smaller than the biological spread of concentrations showing that for those parameters mid-infrared and Raman spectroscopy might supply a valuable tool for quantification. It may be speculated that—similar to the case of glucose, where we have artificially spiked the samples to deliver concentrations of glucose outside the normal, but well within the possible physiological range—the quantification accuracy of protein, HDL, and uric acid may appear more favorable in future studies, using samples which originate from diseased people suffering, e.g., from dyslipidemia or gout.

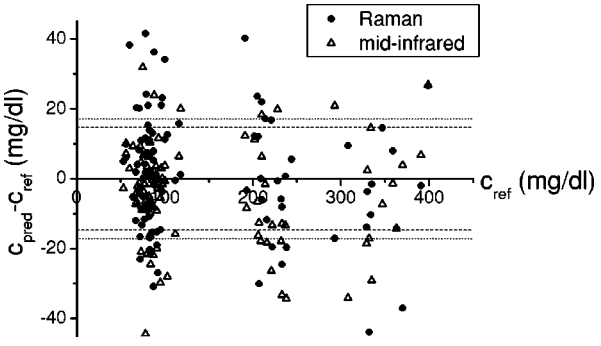

Fig. 5 Difference between the concentration c_{pred} as predicted by Raman (circles) or mid-infrared (triangles) spectroscopy and the concentration c_{ref} as determined by the laboratory analysis. The dashed (dotted) lines indicate the values of ±RMSEP for the mid-infrared (Raman) spectroscopic data.

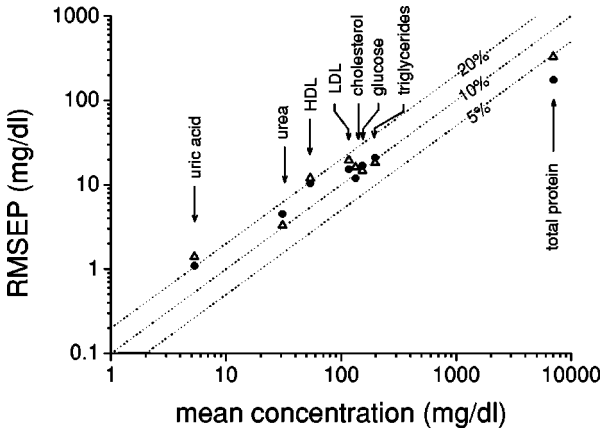

Fig. 6 Root-mean-square error of prediction as a function of mean analyte concentration in the case of Raman (circles) and mid-infrared (triangles) spectroscopy. The dotted lines indicate relative errors of prediction (%RMSEP) of 5%, 10% and 20%.

Table 5 Results of the multivariate analysis of mid-infrared and Raman spectra. The given values are the root-mean-square errors of prediction (RMSEP) of a validation using N_{val} independent samples. As an exception the results reported in Refs. 6 and 8 were obtained using leave-one-out (LOO) validation and the values are therefore marked with an asterisk. For the study reported in this manuscript $N_{val}=99$ samples were subjected to the validation process for both methods. (Note that in the case of infrared spectroscopy one sample was excluded from the evaluation due to unusually low absorbances in all three repetitions of the pipetting.) $N_{tot}=247$ is the total number of samples used in the study. HDL and LDL denote the high and low density lipoprotein fraction of cholesterol, respectively. All concentrations are given in mg/dl.

Reference	N_{tot}	N_{val}	Total protein	Triglycerides	Cholesterol	HDL	LDL	Glucose	Urea	Uric acid
					Mid-infrared					
This study	247	99	328	18.1	16.1	11.9	19.4	14.7	3.3	1.4
[a]	300	100	280	20.1	10.8	…	…	7.4	6.6	2.4
[b]	300	100	310	23.6	11.2	…	…	27	7.2	…
[c]	90	30	…	30.6	14.7	12.0	13.5	…	…	…
[d]	122	24	…	13	15	…	…	16	…	…
[e]	306	(LOO)	240*	16.6*	11.3*	…	…	9.5*	2.0*	…
					Raman					
This study	247	99	176	20.7	11.5	11.0	15.7	17.1	4.4	1.1
[f]	60	18–24	71	…	10.4	…	…	…	…	…
[g]	66	(LOO)	190*	29*	12*	…	…	26*	3.8*	…

[a] Reference 24.
[b] Reference 25.
[c] Reference 26.
[d] Reference 27.
[e] Reference 6.
[f] Reference 18.
[g] Reference 8.

The accuracy of present-day laboratory testing for the parameters investigated is still significantly better than the spectroscopic results. Even more so, the presented results of spectroscopy may be perceived as overoptimistic since, e.g., long-term drifts and instrument-to-instrument variations were purposely avoided in our study. Thus, one might be tempted to conclude that the quantitative analysis of serum based on vibrational spectroscopy cannot compete with present day laboratory diagnostics. However, spectroscopy has the advantage that only one measurement is needed in order to quantify all the shown parameters simultaneously. Furthermore, no reagents are needed for the analysis, thus eliminating reagent costs and reducing logistic efforts. In cases for which moderate accuracy is permissible, vibrational spectroscopy might open the path towards a less expensive and more rapid analysis with the additional benefit of requiring small sample volumes.

Acknowledgments

The work presented in this manuscript has benefited substantially from prior work on mid-infrared spectroscopy at Roche Diagnostics and we thank R. Mischler and G. Werner for sharing their experience. This work would not have been possible without the enthusiasm and expertise of A. Orosz and F. Reichert.

References

1. A. Mahadevan-Jansen, M. G. Sowa, G. J. Puppels, Z. Gyczynski, T. Vo-Dinh, and J. R. Lakowicz, (Eds.) ''Biomedical vibrational spectroscopy and biohazard detection technologies,'' *Proc. SPIE* **5321**, Bellingham, WA (2004).
2. H.-U. Gremlich and B. Yan (Eds.), *Infrared and Raman Spectroscopy of Biological Materials*, Dekker, New York (2001).
3. J. M. Chalmers and P. R. Griffiths (Eds.), *Handbook of Vibrational Spectroscopy*, Vol. 5, Wiley, Chichester (2002).
4. D. Naumann, ''FT-infrared and FT-Raman spectroscopy in biomedical research,'' *Appl. Spectrosc. Rev.* **36**, 238–298 (2001).
5. W. Petrich, ''Mid-infrared and Raman spectroscopy for medical diagnostics,'' *Appl. Spectrosc. Rev.* **36**, 181–237 (2001).
6. G. Werner, D. Boecker, H.-P. Haar, H. J. Kuhr, and R. Mischler, ''Multicomponent assay for blood substrates in human sera and haemolysed blood by mid-infrared spectroscopy,'' in *Infrared Spectroscopy: New Tool in Medicine*, H. Mantsch and M. Jackson (Eds.), *Proc. SPIE* **3257**, 91–100 (1998).
7. S. Low-Ying, R. A. Shaw, M. Leroux, and H. H. Mantsch, ''Quantification of glucose and urea in whole blood by mid-infrared spectroscopy of dried films,'' *Vib. Spectrosc.* **28**, 111–116 (2002).
8. A. J. Berger, T.-W. Koo, I. Itzkan, G. Horowitz, and M. S. Feld, ''Multicomponent blood analysis by near-infrared Raman spectroscopy,'' *Appl. Opt.* **38**, 2916–2926 (1999).
9. A. M. K. Enejder, T.-W. Koo, J. Oh, M. Hunter, S. Sasic, and M. S. Feld, ''Blood analysis by Raman spectroscopy,'' *Opt. Lett.* **27**, 2004–2006 (2002).
10. H. M. Heise, R. Marbach, T. Koschinsky, and F. A. Gries, ''Multicomponent assay for blood substrates in human plasma by mid-infrared spectroscopy and its evaluation for clinical analysis,'' *Appl. Spectrosc.* **48**, 85–95 (1994).

11. H. M. Heise and A. Bittner, ''Multivariate calibration for near-infrared spectroscopic assays of blood substrates in human plasma based on variable selection using PLS-regression vector choices,'' *Fresenius' J. Anal. Chem.* **362**, 141–147 (1998).
12. E. Diessel, S. Willmann, P. Kamphaus, R. Kurte, U. Damm, and H. M. Heise, ''Glucose quantification in dried-down nanoliter samples using mid-infrared attenuated total reflection spectroscopy,'' *Appl. Spectrosc.* **58**, 442–450 (2004).
13. G. Deleris and C. Petibois, ''Application of FT-IR spectroscopy to plasma contents analysis and monitoring,'' *Vib. Spectrosc.* **32**, 129–136 (2003).
14. R. Vonach, J. Buschmann, R. Falkowski, R. Schindler, B. Lendl, and R. Kellner, ''Application of mid-infrared transmission spectrometry to the direct determination of glucose in whole blood,'' *Appl. Spectrosc.* **52**, 820–822 (1998).
15. K. Hebestreit, T. Beyer, A. Lambrecht, R. Mischler, M. Schoemaker, and W. Petrich, ''Infrared spectroscopy of glucose solutions using quantum cascade lasers,'' in *SPIE Technical Summary Digest (BiOS 2004, 5321–31)*, p. 116, SPIE, Bellingham, WA (2004).
16. S. Schaden, M. Haberkorn, J. Frank, J. R. Baena, and B. Lendl, ''Direct determination of carbon dioxide in aqueous solution using mid-infrared quantum cascade lasers,'' *Appl. Spectrosc.* **58**, 667–670 (2004).
17. K. H. Hazen, M. A. Arnold, and G. W. Small, ''Measurement of glucose and other analytes in undiluted human serum with near-infrared transmission spectroscopy,'' *Anal. Chim. Acta* **371**, 255–267 (1998).
18. J. Y. Qu, B. C. Wilson, and D. Suria, ''Concentration measurements of multiple analytes in human sera by near-infrared Raman spectroscopy,'' *Appl. Opt.* **38**, 5491–5497 (1999).
19. D. Rohleder, W. Kiefer, and W. Petrich, ''Raman spectroscopy of serum and serum ultrafiltrate,'' *Analyst (Cambridge, U.K.)* **129**, 906–991 (2004).
20. C. R. Yonzon, C. L. Haynes, X. Zhang, J. T. Walsh, Jr., and R. P. Van Duyne, ''A glucose biosensor based on surface-enhanced Raman scattering: improved partition layer, temporal stability, reversibility and resistance to serum protein interference,'' *Anal. Chem.* **76**, 78–85 (2004).
21. J. Moecks, G. Kocherscheidt, W. Köhler, and W. Petrich, ''Progress in diagnostic pattern recognition,'' *Proc. SPIE* **5321**, 117–123 (2004).
22. J. Moecks, D. Rohleder, and W. Petrich, German patent application (pending).
23. C. A. Lieber and A. Mahadevan-Jansen, ''Automated method for subtraction of fluorescence from biological Raman spectra,'' *Appl. Spectrosc.* **57**, 1363–1367 (2003).
24. R. A. Shaw, S. Kotowich, M. Leroux, and H. H. Mantsch, ''Multianalyte serum analysis using mid-infrared spectroscopy,'' *Ann. Clin. Biochem.* **35**, 624–632 (1998).
25. R. A. Shaw and H. H. Mantsch, ''Multianalyte serum assay from mid-infrared spectra of dry film on glass slides,'' *Appl. Spectrosc.* **54**, 885–889 (2000).
26. K.-Z. Liu, R. A. Shaw, A. Man, T. C. Dembinski, and H. H. Mantsch, ''Reagent-free, simultaneous determination of serum cholesterol in HDL and LDL by infrared spectroscopy,'' *Cin. Chem.* **48**, 499–506 (2002).
27. W. Petrich, B. Dolenko, J. Früh, M. Ganz, H. Greger, S. Jacob, F. Keller, A. E. Nikulin, M. Otto, O. Quarder, R. L. Somorjai, A. Staib, G. Werner, and H. Wielinger, ''Disease pattern recognition in infrared spectra of human sera with diabetes mellitus as an example,'' *Appl. Opt.* **39**, 3372–3379 (2000).