



ELSEVIER

Analytica Chimica Acta 371 (1998) 255–267

**ANALYTICA  
CHIMICA  
ACTA**

# Measurement of glucose and other analytes in undiluted human serum with near-infrared transmission spectroscopy

Kevin H. Hazen<sup>1,a</sup>, Mark A. Arnold<sup>a,\*</sup>, Gary W. Small<sup>b</sup>

<sup>a</sup>Department of Chemistry, Iowa Advanced Technology Laboratories, University of Iowa, Iowa City, IA 52242, USA

<sup>b</sup>Center for Intelligent Chemical Instrumentation, Department of Chemistry, Clippinger Laboratories, Ohio University, Athens, OH 45701–2979, USA

Received 14 January 1998; received in revised form 26 April 1998; accepted 28 April 1998

## Abstract

Near-infrared calibration models are described for the measurement of total protein, albumin protein, globulin protein, triglycerides, cholesterol, urea, glucose, and lactate. Spectra are collected in triplicate over the 5000–4000  $\text{cm}^{-1}$  spectral range with a 2.5 mm optical path length for 242 undiluted human serum samples. Calibration models are generated for each analyte by performing partial least-squares (PLS) regression on raw and digitally filtered spectra. Models are optimized individually for each analyte by considering spectral range, number of model factors and width/position of a Gaussian shaped filter response function for a digital Fourier filter. Accurate measurements are demonstrated for each analyte except lactate which is below the detection limit under our experimental conditions. Standard error of prediction (SEP) for glucose is 23.3 mg/dl (1.29 mM). Relevance and stability of the glucose calibration model are examined by evaluating the accuracy of glucose predictions from 50 human serum samples collected on a modified spectrometer nineteen months after the calibration data were collected. In addition, these 50 subsequent samples were treated in a blind manner by withholding their glucose values until all predictions were complete. Results indicate a slight positive bias in the predictions corresponding to a minor instability in the model. This instability is likely due to changes in the spectrometer hardware. Nevertheless, the strong correlation between predicted and actual glucose levels in these blind samples strongly suggests that this calibration model is based on information particular to glucose. © 1998 Elsevier Science B.V. All rights reserved.

**Keywords:** Near-infrared spectroscopy; NIR; PLS; Glucose sensing

## 1. Introduction

Near-IR spectroscopy holds great promise for clinical chemistry measurements on the basis of its

potential for reagent-less, nondestructive, and noninvasive measurements [1–3]. Conceptually, the measurement involves collecting near-IR spectra from raw clinical samples and then building a multivariate calibration model that relates spectral variations to the desired analytical information. The ability to collect meaningful data without diluting the sample or adding reagents minimizes sample handling. The nondestructive nature of the measurement permits

\*Corresponding author. Tel.: 001 319 335 1368; fax: 001 319 353 1115; e-mail: mark-arnold@uiowa.edu

<sup>1</sup>Present address: Instrumentation Metrics, Inc., 2085 Technology Circle, Suite 102, Tempe, AZ 85284, USA.

subsequent analyses with the same sample aliquot, thereby reducing sample volume requirements. If the method can be successfully implemented noninvasively [1,4,5], the need to collect a sample will be eliminated altogether and the ability to monitor *in vivo* levels of selected chemical species in a continuous manner will be available to the clinician.

Several reports describe clinical chemistry measurements with vibrational spectroscopy [6–15]. The work published by Hall and Pollard [3,16] focuses directly on the analysis of human serum by near-IR spectroscopy. In this work, near-IR spectra were collected over the  $25\,000\text{--}4000\text{ cm}^{-1}$  (400–2500 nm) spectral range from 0.5 mm thick samples of human serum. Multivariate linear regression (MLR) methods were used to build calibration models for total protein, serum albumin and globulin proteins and partial least-squares (PLS) regression was used to construct models for triglycerides, urea, and glucose. Spectral information corresponding to both the overtone and combination regions were incorporated into these PLS calibration models. Specifically, the spectral ranges used by Hall and Pollard were  $7547\text{--}5555$  and  $4914\text{--}4211\text{ cm}^{-1}$  for triglycerides, urea and glucose models. Their reported standard error of calibration (SEC) values are reasonable for all these analytes compared to their standard reference methods. Furthermore, models for proteins, triglycerides, and urea were validated by computing standard error of prediction (SEP) values from an independent set of sample spectra. No SEP values were reported for glucose, however [16].

We are interested in assessing the analytical utility of near-IR spectroscopy for clinical and biochemical measurements [17–25]. Our efforts have focused exclusively on the first combination region of the near-IR spectrum which extends from  $4850$  to  $4200\text{ cm}^{-1}$  in aqueous solutions. Absorption features in this spectral region correspond to combinations of vibrational stretching and bending frequencies associated with CH and NH molecular features. Absorptivities are small for these combination bands relative to the fundamental transitions. In addition, the ambiguity of CH and NH groups in biochemical molecules results in significant spectral overlap, which complicates interpretation and quantitation.

In this paper, we expand the work of Hall and Pollard by establishing the utility of near-IR clinical

measurements based solely on the combination spectral range. PLS calibration models are generated for total protein, albumin protein, globulin protein, cholesterol, triglycerides, urea, glucose and lactate. Acceptable prediction ability is demonstrated for all these analytes, except lactate, which is present at concentrations below the detection limit of the method. In addition, a blind sample protocol is used to validate the prediction ability for glucose. Lastly, calibration model stability is tested by analyzing a set of serum spectra collected several months after the training spectral data set.

## 2. Experimental

### 2.1. Instrumentation

Spectra were collected with a modified Nicolet 740 Fourier transform spectrometer [26]. This spectrometer was equipped with a 250 W tungsten-halogen lamp (Gilway Technical Lamps, Waltham, MA), a spherical back-reflecting mirror on the lamp housing, a calcium fluoride beam splitter, and a cryogenically cooled indium antimonide (InSb) detector. A multi-layer optical interference filter (Barr and Associates, Westford, MA) was used to isolate the  $5000\text{--}4000\text{ cm}^{-1}$  spectral range. Samples were held in a 2.5 mm path length, thermostatted sample holder. The sample temperature was maintained at  $37.0 (\pm 0.1)^\circ\text{C}$  with a VWR model 1140 refrigerated temperature bath (VWR Scientific, Chicago, IL). Temperatures were measured by inserting a 1/32 in diameter copper-constantan thermocouple probe (Omega, Stamford, CT) directly into the sample solution and reading values from an Omega model 670 digital meter.

### 2.2. Samples and reagents

Human serum samples were obtained from the Department of Pathology at the University of Iowa Hospitals and Clinics. Serum was prepared by centrifuging coagulated blood into serum separator tubes without fibrinogen. Samples were collected randomly from the general hospital population with the exception that special effort was made to collect a reasonable number of hyperglycemic samples.

Each sample was analyzed by the Clinical Chemistry Laboratory at the University of Iowa Hospitals and Clinics and their results were used as reference values for all analytes except glucose. Reference glucose concentrations were measured with a YSI model 2300 Stat Plus analyzer at the same time the spectrum was collected. This analyzer provides glucose values with a standard prediction error of 0.25 mM [27]. Samples were collected over a five week period and each sample was frozen to preserve integrity. Samples were thawed slowly to room temperature and then gently heated to  $37.0 \pm 0.1^\circ\text{C}$  just before collecting spectra.

All standard solutions were prepared by weighing suitable amounts of reagent grade material and diluting with a stock working buffer. The working buffer was a pH 7.3, 0.1 M phosphate buffer prepared with reagent grade water.

### 2.3. Procedures

Spectral information was collected as 256 co-added, double-sided interferograms with 16 384 points. Interferograms were converted to single-beam spectra by using appropriate functions within the Nicolet SX-FTIR ver 4.4 software package. Point spacing for the resulting spectra was  $1.9\text{ cm}^{-1}$ . Single-beam spectra were transferred to a Silicon Graphics Iris Indigo R4000 workstation (Silicon Graphics, Mountain View, CA) for further processing. All spectral processing software on the Silicon Graphics computer was implemented in Fortran 77 and Fourier filtering calculations used subroutines from the IMSL software library (IMSL, Houston, TX).

Calibration models were constructed based on PLS regression with and without digital Fourier filtering. Specific procedures used to build and validate these models are detailed elsewhere [19,20,22,24].

Single-beam spectra were collected in triplicate for each of 242 individual human serum samples. Table 1 summarizes the distribution of analytes within this set of samples. Spectra for these original 242 samples were divided into calibration, monitoring and prediction data sets. For this procedure, samples were selected randomly and all replicate spectra for a given sample were moved together. The final distribution resulted in 486 spectra (162 samples) in the calibration set and 120 spectra (40 samples) in each of the

Table 1  
Analyte distribution in human serum samples<sup>a</sup>

Analyte	Mean	Standard deviation	High	Low
Total protein	6980	640	8500	5400
Serum albumin	4210	540	5500	2600
Globulin protein	2760	470	4200	1600
Triglycerides	213	141	840	40
Cholesterol	211	52	384	86
Urea	19.2	14.0	143.0	4.0
Glucose	202	113	574	57
Lactate	20.5	7.9	50.8	7.8

<sup>a</sup>All concentrations in mg/dl.

monitoring and prediction data sets. As described before [19], the monitoring set was used to identify the optimum mean position and standard deviation width for the Gaussian shaped bandpass filter used in the Fourier filtering procedure.

A second group of 56 serum samples was collected several months later. This group was used to test the prediction accuracy of the original calibration models for glucose. Six of these samples were used to evaluate potential success before undergoing the larger experiment. For the remaining 50 samples, a blind protocol was used by leaving the concentrations of glucose for these samples with the head of the Clinical Chemistry Laboratory until the analysis was complete and predictions were established. Again, triplicate spectra were collected for each sample.

### 3. Results and discussion

Protein and water are the principal components of human serum. The near-IR spectrum of human serum is dominated by the strong absorbances of these two components. The major influence of water has been eliminated in this work by using background spectra of a 0.1 M phosphate buffer. Protein features dominate the resulting absorbance spectra. This point is illustrated in Fig. 1 which shows absorbance spectra for a typical serum sample superimposed on absorbance spectra for albumin and globulin proteins. The concentrations of triglycerides, cholesterol, urea, glucose, and lactate are well below the concentrations of these proteins which establish the challenge in selectivity extracting analyte specific spectral information super-

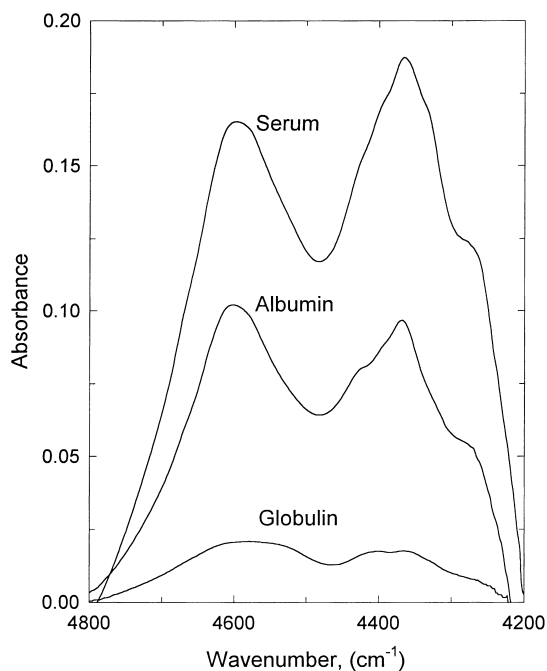


Fig. 1. Sample absorbance spectra for human serum, albumin protein and globulin protein.

imposed on the stronger and overlapping absorbances from these more concentrated substances.

### 3.1. Analyte co-variance

As noted before for PLS [5,22], independent calibration models require no co-variance between multiple components within the sample matrix. The concentration of each analyte must vary independently within the data set, so spectral features for one analyte will not be falsely incorporated into the calibration

model for another analyte. To test for co-variance within our data set of human spectra, correlation plots were constructed for each combination of analytes. A linear regression analysis of these plots provided the correlation coefficients summarized in Table 2. The largest correlations are noted for albumin and globulin proteins with total protein which is understandable considering total protein is composed of these two classes of protein. In fact, reference values for globular protein are computed as the difference between measured values for total and albumin protein. Nevertheless, correlations are small between proteins and correlations are virtually zero for all other species.

### 3.2. Optical path length

Selection of a single path length for the measurement of numerous analytes can be difficult given differences in location and size of pertinent analyte absorption bands. The goal, of course, is to maximize the signal-to-noise ratio (SNR) by using long path lengths to increase the analyte-dependent absorbance signal under conditions of low spectral noise.

Spectral noise is frequency dependent, however, because of the absorption characteristics of water. Water possesses two broad absorption bands centered at 3800 and 5200  $\text{cm}^{-1}$ . These bands absorb significant radiation throughout the 5000–4000  $\text{cm}^{-1}$  spectral region. A minimum absorption is observed at 4516  $\text{cm}^{-1}$  for 37°C water. Absorbance at 4516  $\text{cm}^{-1}$  is approximately 1 absorbance unit (AU) per mm path length of water and absorbance values increase dramatically at both higher and lower wave numbers. Operation of the spectrometer in a detector noise limited condition results in different noise levels across the spectrum as the optical throughput is

Table 2  
Test for co-variance between analytes<sup>a</sup>

	Albumin	Globulin	Triglycer	Cholest	Urea	Glucose	Lactate
Total protein	0.459425	0.356144	0.001828	0.044357	0.067874	0.030887	0.003382
Albumin		0.034696	0.000825	0.056246	0.061404	0.047181	0.003443
Globulin			0.007816	0.000492	0.005921	0.000024	0.000190
Triglycer				0.180583	0.008606	0.019033	0.039299
Cholest					0.007876	0.009349	0.003140
Urea						0.045171	0.003335
Glucose							0.054451

<sup>a</sup>r-Square values from linear least-squares regression.

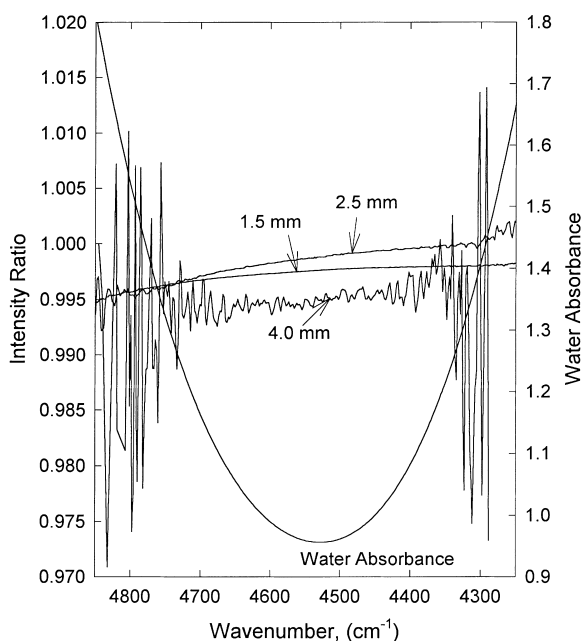


Fig. 2. Representative 100% lines for indicated optical path lengths for human serum samples. An absorbance spectrum of water is superimposed to illustrate optical throughput.

limited by the absorption of water. This effect is illustrated in Fig. 2 which shows 100% lines for different thicknesses of human serum superimposed on an absorbance spectrum of a 1 mm thick sample of water. For a 2.5 mm path length, the root-mean-square (RMS) noise computed across the entire spectral range (5000–4000  $\text{cm}^{-1}$ ) is 149 327  $\mu\text{AU}$  compared to only 523  $\mu\text{AU}$  when the range is reduced to 4850–4250  $\text{cm}^{-1}$ . Noise levels are further reduced by narrowing the spectral range closer to the water absorption minimum. The RMS noise is only 37  $\mu\text{AU}$  for the 4457–4354  $\text{cm}^{-1}$  spectral range which corresponds to the central absorption feature of glucose [17,20,22]. For absorption features with equivalent absorptivities, the SNR values will be greatest for spectral bands closest to 4516  $\text{cm}^{-1}$ .

Spectral noise also depends on sample thickness. As shown in Fig. 2, thicker samples result in lower optical throughput and greater spectral noise. The RMS noise increases exponentially with path length. For example, RMS noises for the 4457–4354  $\text{cm}^{-1}$  spectral range are 12.1, 21.2, 37.0, 92.1, 363, and 1056  $\mu\text{AU}$  for path lengths of 1.5, 2.0, 2.5, 3.0, 3.5, and 4.0 mm, respec-

tively. Clearly a tradeoff is needed in terms of optical path length. The path length must be sufficiently long to maximize analyte-dependent absorbances yet short enough to provide low spectral noise.

A 2.5 mm path length was selected for this study on the basis of the primary objective of measuring glucose. Our previous work indicates that a RMS noise level of 20  $\mu\text{AU}$  was sufficient to measure glucose in 1 mm thick aqueous solutions [20]. As noted above for the 4457–4354  $\text{cm}^{-1}$  spectral range, the RMS noise level is 37.0  $\mu\text{AU}$  for a 2.5 mm thickness of serum. This value corresponds to 14.8  $\mu\text{AU}/\text{mm}$  which, in the absence of chemical interferences, should provide sufficient SNR for the measurement.

### 3.3. Protein measurements

Albumin and globulin proteins are major components of human serum and, as noted before, the predominant spectral features of serum correspond to these proteins. Hence, calibration models for protein should be relatively straightforward. An interesting question, however, is the ability to differentiate albumin and globulin proteins. As illustrated in Fig. 1, their absorbance spectra are similar in terms of number and position of absorption features. The only differences correspond to relative magnitude and shape. We have previously established the feasibility of constructing selective calibration models based on slight spectral differences [28]. In this previous work, selective calibration models were developed for glutamine and asparagine by using PLS regression methods. The results presented below demonstrate the successful differentiation of albumin and globulin proteins by PLS calibration models.

Independent calibration models were constructed and evaluated for total protein, albumin protein and globulin protein. The entire spectral range from 4850 to 4250  $\text{cm}^{-1}$  was used for the total protein model, while a series of spectral ranges were tested for albumin and globulin proteins. Spectral ranges tested for albumin were 4850–4250, 4700–4300, 4700–4475, and 4475–4300  $\text{cm}^{-1}$ , while tested ranges for globulin protein were 4850–4250, 4700–4325, 4700–4470, and 4470–4325  $\text{cm}^{-1}$ . In both cases, these ranges correspond to the whole spectral range, a slightly narrower range that encompasses both absorbance features, and ranges that isolate the high and

Table 3  
Optimum calibration models for protein

Protein type	Spectral range (cm <sup>-1</sup> )	Spectral features	Digital filter mean ( <i>f</i> )	Digital filter standard deviation ( <i>f</i> )	Number of PLS factors	SEC (mg/dl)	SEP (mg/dl)
Total	4850–4250	Both bands	—	—	10	0.21	0.23
	4850–4250	Both bands	0.02	0.0085	7	0.22	0.24
Albumin	4850–4250	Both bands	—	—	7	0.20	0.20
	4850–4250	Both bands	0.052	0.0130	4	0.20	0.19
	4700–4300	Both bands	—	—	10	0.19	0.19
	4700–4300	Both bands	0.017	0.003	3	0.20	0.21
	4470–4475	High freq	—	—	10	0.18	0.19
Globulin	4475–4300	Low freq	—	—	12	0.19	0.20
	4850–4250	Both bands	—	—	12	0.21	0.23
	4850–4250	Both bands	0.039	0.0120	10	0.20	0.22
	4700–4325	Both bands	—	—	10	0.20	0.21
	4700–4325	Both bands	0.029	0.0055	4	0.22	0.21
	4700–4470	High freq	—	—	9	0.21	0.21
	4470–4325	Low freq	—	—	8	0.26	0.27

low wave number bands, respectively (see Fig. 1). For each analyte, and then for each spectral range, a series of PLS calibration models were generated by using 1–20 factors. Model performance was judged by evaluating values for the SEC and SEP.

Calibration models for total, albumin, and globulin protein levels are essentially equivalent, regardless of protein type or spectral range. Table 3 summarizes the pertinent results. Inspection of these values reveals errors of calibration and prediction around 0.2 g/l which matches the reference method for protein [29]. In addition, equivalent performance is obtained when a digital Fourier filtering step is implemented prior to the PLS regression. As indicated in Table 3, standard errors are still 0.2 g/l. As we have noticed with other data sets, the incorporation of a digital filtering step decreases the number of latent variables required to achieve the same level of performance [18,19]. For the five spectral regions tested, an average of 4.2 fewer factors is required to obtain the minimum SEP. Fourier filtering is known to effectively reduce both baseline variations and high frequency spectral noise [19] which indicates that some of the factors in the PLS models with raw spectra are needed to compensate for these types of spectral features.

Several critical points are apparent by plotting SEC and SEP as a function of the number of factors used in the PLS model. Such a plot is presented in Fig. 3 for total, albumin, and globulin protein models computed

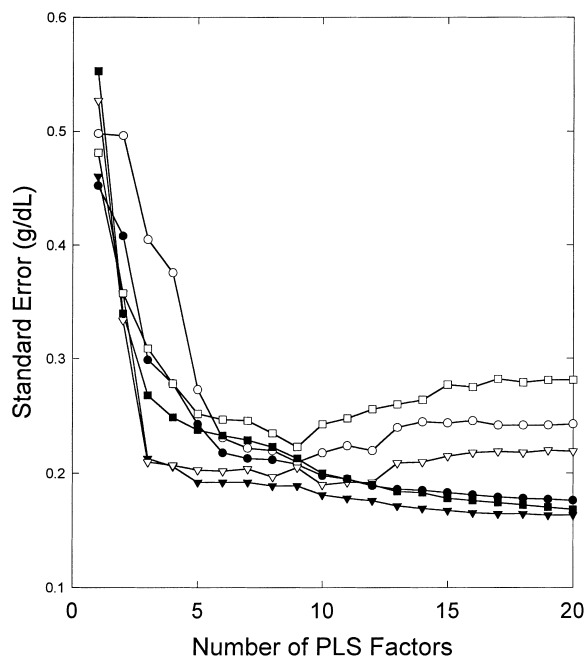


Fig. 3. Effect of number of PLS factors on SEC (closed) and SEP (open) for total protein (squares), albumin protein (triangles) and globulin protein (circles).

over the 4850–4250 cm<sup>-1</sup> spectral range. The first point of interest is a minimum in SEP. Initially, both SEP and SEC drop as the early factors account for much of the relevant spectral variation. The relatively

large decrease in standard errors indicates that a majority of the protein-specific information is incorporated into the first three or four factors. The extent of decline diminishes beyond this point as only subtle spectral features are incorporated into the models. A second point is the continual decline in SEC which indicates overmodeling. Both the minimum in SEP and continual decline in SEC are consistent with earlier observations with less complicated data sets [18,23].

A final interesting point from Fig. 3 is the order in which the standard errors decline. The sharpest drop corresponds to albumin protein, followed by total protein and then globulin protein. This order is consistent with the relative magnitude of absorbances for albumin and globulin proteins. The higher concentration of albumin protein results in greater absorbances which dominate the spectra and, thereby, are easier to incorporate into the calibration model. The smaller globulin absorbances are more difficult to extract in the presence of the overshadowing albumin absorbance features. The first two factors for globulin models must account for spectral variations due to albumin; hence SEP for globulin protein is largely unaffected. Eventually, the same model performance is obtained but with more factors. Models for total protein require information about both protein types which places it between the curves for albumin and globulin taken separately.

### 3.4. Other analytes

High levels of serum protein create a challenge for measuring other serum components that are present at lower concentrations and, hence, do not generate such strong near-IR absorbances. The ability to establish accurate calibration models for triglycerides, cholesterol, urea, glucose and lactate has been assessed. Functional models are possible for each of these analytes, except lactate which is present at concentrations below the detection limit of the method.

The key to successful differentiation of these numerous analytes lies in their spectral differences. Fig. 4 shows a series of normalized absorbance spectra for these analytes. Although the absorption features overlap extensively, each spectrum is unique. These spectra are also different compared to those for albumin and globulin proteins as shown in Fig. 1.

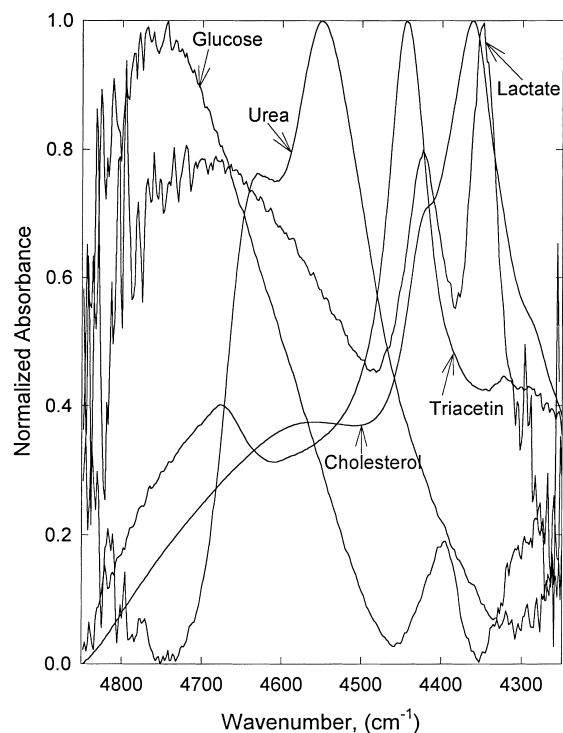


Fig. 4. Normalized absorbance spectra for 300 mg/dl triacetin, 200 mg/dl cholesterol, 139 mg/dl urea, 180 mg/dl glucose, and 74.8 mg/dl lactate.

A series of PLS calibration models was constructed and evaluated. In each case, models were generated with 1–30 factors. For each of these models, all combinations of mean centering and spectral normalization were tested which resulted in four models with each factor (no mean centering or normalization; mean centering without normalization; normalization without mean centering; mean centering with normalization). The spectral ranges tested for each analyte included the whole 4850–4250  $\text{cm}^{-1}$  range and several more narrowly defined ranges which were selected on the basis of the positions of known absorption bands for a particular analyte (see Fig. 4). Additional ranges used for cholesterol, as an example, are 4500–4250 and 4450–4300  $\text{cm}^{-1}$  which encompass the large absorbance features of cholesterol located in the low frequency region of the spectrum. Each spectral range tested is listed in Tables 4 and 5 which summarize conditions and results from the best PLS models with raw and Fourier filtered spectra, respec-

Table 4  
Performance of PLS models with raw spectra

Analyte	Spectral range (cm <sup>-1</sup> )	Factors	mc <sup>a</sup>	nor <sup>b</sup>	SEC (mg/dl)	SEP (mg/dl)	MPEP <sup>c</sup> (%)
Triglycerides	4850–4250	15	n	y	12.2	13.9	6.8
	4800–4625	6	y	n	52.2	49.4	25.5
	4675–4500	10	n	n	42.7	39.9	23.1
	4525–4375	11	n	n	41.6	41.3	20.0
Cholesterol	4850–4250	18	n	y	8.4	12.6	5.3
	4500–4250	10	n	n	11.9	12.5	5.0
	4450–4300	15	y	y	9.9	13.8	5.4
Urea	4850–4250	29	n	y	0.84	1.3	5.9
	4800–4300	17	n	y	0.95	1.2	5.2
Glucose	4850–4250	11	y	n	25.4	24.5	13.4
	4811–4457	11	y	y	26.8	29.4	20.1
	4457–4354	8	n	y	44.1	37.1	20.8
	4380–4227	16	y	y	58.5	54.9	31.4
Lactate	4850–4250	22	n	y	4.5	6.9	29.5
	4575–4380	2	n	n	7.6	9.4	42.3
	4380–4310	1	y	n	7.6	9.5	44.0

<sup>a</sup>mc=mean centered.

<sup>b</sup>nor=normalized.

<sup>c</sup>MPEP=mean percent error of prediction.

Table 5  
Performance of PLS models with Fourier filtered spectra

Analyte	Spectral range (cm <sup>-1</sup> )	Filter mean ( <i>f</i> )	Filter standard deviation ( <i>f</i> )	Factors	mc <sup>a</sup>	nor <sup>b</sup>	SEC (mg/dl)	SEP (mg/dl)	MPEP <sup>c</sup> (%)
Triglycerides	4850–4250	0.051	0.0176	13	y	y	11.9	10.4	5.4
	4700–4470	0.045	0.0040	12	y	y	20.2	16.8	8.9
Cholesterol	4850–4250	0.067	0.0175	13	y	n	11.4	12.1	4.9
	4500–4250	0.049	0.0150	10	n	n	12.3	12.8	5.2
Urea	4850–4250	0.034	0.0150	12	y	y	1.4	1.3	6.0
	4800–4300	0.034	0.0105	12	n	n	1.1	1.2	5.4
Glucose	4850–4250	0.039	0.0140	13	n	y	21.6	23.3	15.1
	4811–4457	0.002	0.0055	17	y	y	25.8	26.7	16.5
	4457–4354	0.008	0.0250	9	n	y	41.1	31.7	21.5
	4357–4227	0.014	0.0070	12	n	y	55.4	56.9	26.7
Lactate	4850–4250	0.042	0.0095	24	n	n	5.3	6.1	24.1

<sup>a</sup>mc=mean centered.

<sup>b</sup>nor=normalized.

<sup>c</sup>MPEP=mean percent error of prediction.

tively. A comparison of values for each analyte reveals a consistent pattern where models with digital filtering provide similar prediction performance with fewer factors and less sensitivity to spectral range.

Examples of concentration correlation plots for triglycerides, cholesterol, urea and lactate are presented in Fig. 5. The triglyceride plot in Fig. 5 corre-

sponds to a model over the 4850–4250 cm<sup>-1</sup> range with 13 PLS factors and a Gaussian-shaped Fourier filter defined by a bandpass position of 0.051*f* and a bandpass width of 0.0176*f*. The width is specified as the standard deviation of the Gaussian function. The filter parameters are specified in digital frequency units (*f*), where the range 0–0.5*f* defines the range

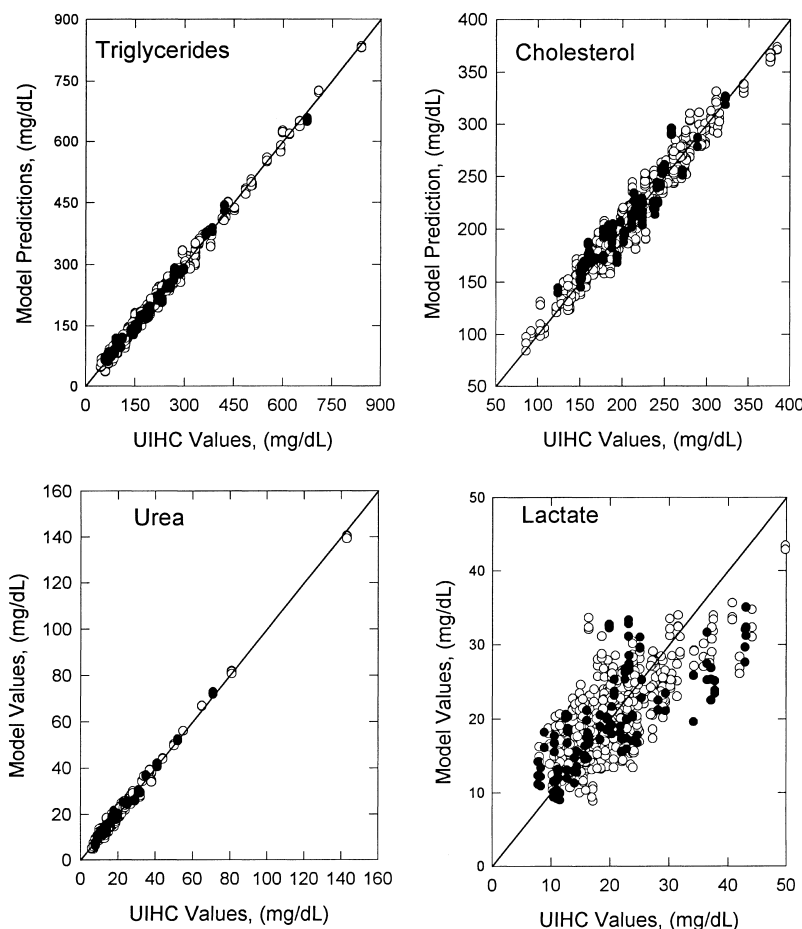


Fig. 5. Representative concentration correlation plots for triglycerides, cholesterol, urea, and lactate showing calibration (open circles) and prediction (closed circles) data points.

of harmonic frequencies comprising the data. The high concentration of triglycerides and their sharp absorption features result in an excellent calibration model with both calibration and prediction data points tightly clustered around the unity line. Although the plot for cholesterol exhibits more scatter, both calibration and prediction points follow the unity line and the scatter of the prediction points is within that of the calibration points. This particular plot for cholesterol corresponds to data from the  $4500\text{--}4250\text{ cm}^{-1}$  model described in Table 5. Similarly, the plot for urea demonstrates excellent correlation between values predicted with the  $4850\text{--}4250\text{ cm}^{-1}$  model described in Table 5 and the urea reference values.

Results for the best lactate model, shown in Fig. 5, indicate poor correlation between the model and the lactate reference values. The mean concentration of lactate in these samples is 20.5 mg/dl (2.28 mM). This low concentration coupled with the lack of an isolated absorption band and low molar absorptivity makes it impossible to measure lactate accurately in this matrix. The SEP for this model is 6.1 mg/dl (0.68 mM) which is similar to the standard deviation of 7.9 mg/dl (0.88 mM) for lactate in our set of serum samples (see Table 1).

Results from the  $4850\text{--}4250\text{ cm}^{-1}$  glucose calibration model in Table 5 are presented in Fig. 6. This concentration correlation plot is superimposed on the

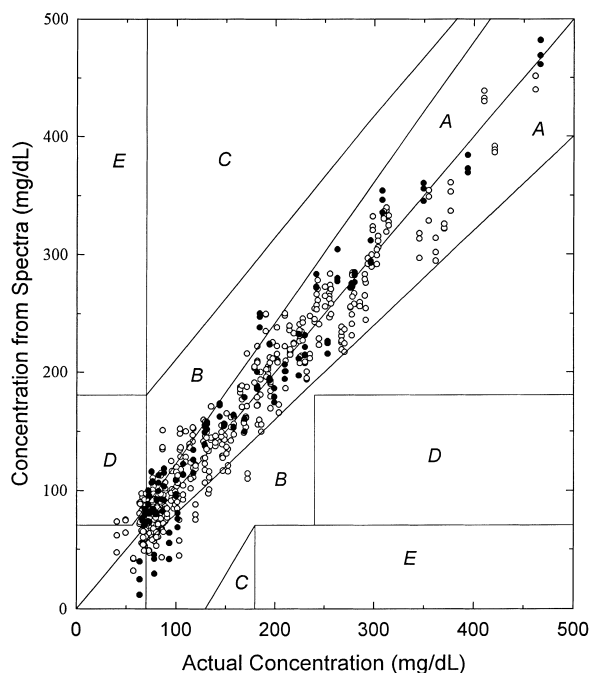


Fig. 6. Serum glucose measurements compared to Clarke error grid for both calibration (open circles) and prediction (closed circles) data points.

Clarke error grid [30]. As is detailed elsewhere, this grid indicates the clinical relevance of measurement accuracy. Points in region A correspond to measurements with sufficient accuracy that the correct clinical decision is reached based on the measured value. Region B corresponds to less accurate measurements with essentially no adverse clinical effects. Region C represents a more dangerous region resulting in clinical decisions that overcorrect acceptable glucose levels. Region D indicates a normal glucose level for individuals that are actually either hypoglycemic or hyperglycemic. This situation is particularly dangerous for hypoglycemic patients. Finally, region E corresponds to the worst situation where the exact opposite clinical decision is made on the basis of faulty analytical results. Regions A and B are considered clinically acceptable while regions C, D, and E are not acceptable.

Accuracy of our serum glucose measurements is borderline clinically acceptable. All values above 200 mg/dl (11.1 mM) fall within region A. A significant fraction of the lower concentration data falls in

region B. Several observations below 65 mg/dl (3.6 mM) are in region D which is clinically unacceptable. Nevertheless, the majority of points reside in region A which attests to the potential value of this methodology. Clearly, the magnitude of the standard errors of calibration and prediction must be reduced in order to meet clinical demands. Lower errors can be achieved by raising the source power which will increase optical throughput, thereby reducing spectral noise.

Although the SEP reported here for glucose is too high for clinical acceptance, the fact that glucose can be measured at all is significant. The sample matrix in this experiment corresponds to 242 individual serum samples. Each sample has a unique chemical composition and contains different levels of multiple near-IR absorbing species. Successful glucose measurements in such a complex and varying chemical matrix suggest that a global calibration model is possible for measurements in undiluted human serum.

### 3.5. Measurement of glucose in blind serum samples

The nature of the PLS regression makes it difficult to verify that the information used to predict glucose levels in serum samples actually originates from the glucose molecule as opposed to some instrument or sample variation that happens to correlate with glucose. Such glucose/system correlations are more likely to occur: (1) when the entire data set is collected during a short period of time (i.e., hundreds of spectra collected over several days); (2) when all spectra are collected under essentially identical experimental conditions (i.e., same instrument settings, same instrument components, same daily laboratory temperature fluctuations, etc.); (3) when the calibration and prediction spectra are collected randomly during the same period. Ideally, accurate glucose predictions can be obtained from spectra collected long after the calibration model is established and under different laboratory conditions. This concept was tested by analyzing glucose predictions computed from a PLS calibration model where the calibration and prediction spectra were collected at different times and under different conditions. In addition, the prediction spectra were handled as blind samples where the glucose values were unknown until after the PLS analysis was complete.

Nineteen months after the spectra described above were collected, a second set of prediction spectra was collected from 56 new serum samples. In the intervening time, the 250 W source was replaced by a new 400 W source, the beam splitter was serviced and realigned, the sample holder was disassembled and the entire instrument was moved to a new laboratory in a different building. The idea was to see how well the calibration model established from the previously collected spectra could predict glucose from spectra collected following these rather dramatic changes in instrument and environment.

Before collecting spectra of the 50 new samples, several old serum samples were thawed and used to set the experimental conditions. The sample compartment was reassembled with the same nominal sample thickness as before and the source power was adjusted until single-beam intensities of these old serum samples matched those collected before. After instrument adjustments, a set of six test samples were obtained and analyzed to give some indication of success before committing the time required to run 50 new samples. Results for the six test samples were encouraging with an SEP of 16.6 mg/dl (0.92 mM) which compares favorably to the original value of 23.3 mg/dl (1.29 mM). Next, a set of 50 new serum samples was then obtained, spectra were collected in triplicate, and glucose values were predicted. As noted in Section 2, glucose concentrations in these new samples were not released from the Clinical Chemistry Laboratory until after predictions were completed.

A calibration model was generated by using all 726 spectra, corresponding to 242 samples, collected before and described above. A numerical optimization procedure based on the use of genetic algorithms was employed to design a digital filter for preprocessing the spectra and a PLS calibration model for use in predicting the glucose concentrations [24]. The resulting model was based on the spectral range from 4850 to 4295  $\text{cm}^{-1}$ , a Fourier filter defined by a filter bandpass position and width of 0.0447 and 0.0172 $f$ , respectively, and a PLS regression model based on 15 latent variables. This model was characterized by an SEC of 23.7 mg/dl (1.32 mM).

Comparison of glucose values predicted for the 50 blind samples from the PLS model and those supplied from the Clinical Chemistry Laboratory revealed an SEP of 52.5 mg/dl (2.91 mM) which is considerably

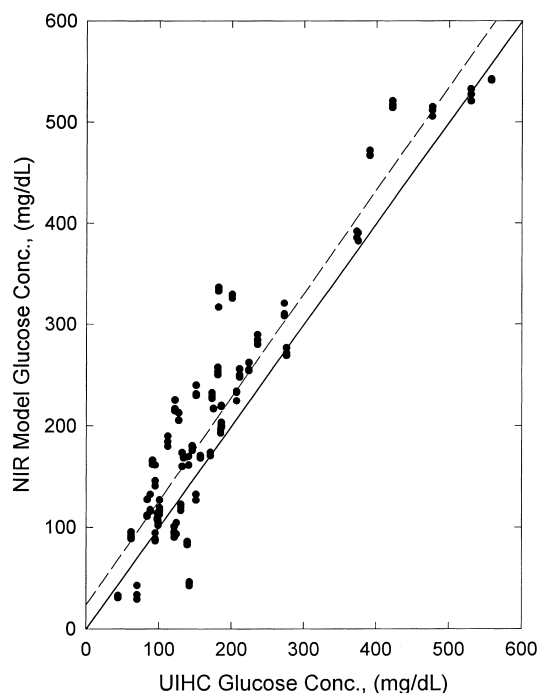


Fig. 7. Prediction correlation plot for blind serum samples. Solid and broken lines represent ideal correlation and regression line, respectively.

greater than the SEC of the model and the SEP computed from the original data set. Inspection of the concentration correlation plot for these data, which is presented in Fig. 7, indicates a clear positive bias in the prediction values. Regression analysis reveals a slope of 1.02, a  $y$ -intercept of 23.7 mg/dl (1.32 mM) and  $r^2$  values of 0.88406. Further analysis shows that this bias is not related to total protein, lipemic index (a measure of sample turbidity), or order analyzed. Systematic errors of this nature were investigated by analyzing the residual plots presented in Fig. 8. Although protein is a major serum component and strong near-IR absorber, it has no effect on prediction accuracy. Likewise, no correlation exists with lipemic index which is interesting given the fact that samples with high lipemic index scatter light and appear cloudy. The lack of interference from this turbidity suggests that the digital filtering step is effective in eliminating this type of baseline variation as suggested in our earlier work [20]. A slight positive correlation is evident between the glucose residuals and the order in which the samples were analyzed. This correlation is

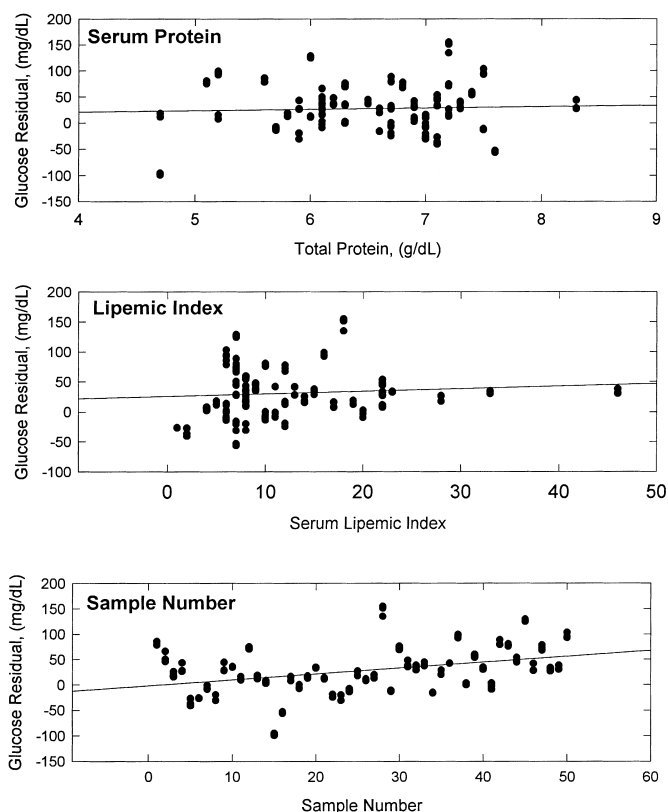


Fig. 8. Residual plots for glucose predictions as functions of total protein, lipemic index and order of analysis. Solid lines indicate linear regression fit to each data set.

small, however ( $r^2=0.14163$ ), and likely corresponds to slight variations in the experimental conditions during the course of the data collection.

The positive bias described above and illustrated in Fig. 7 represents instability in the calibration model related to some physical parameter that was not properly controlled. This level of bias is not surprising given the nature of the experiment, time between calibration and prediction data sets, and modifications in the spectrometer hardware. An off-set correction was applied by subtracting the y-intercept from the linear regression analysis of the correlation plot shown in Fig. 7. The resulting correlation plot (after off-set correction) is presented in Fig. 9 which shows an even distribution of glucose predictions about the unity line. Clearly, the information used for these predictions is not random or chance correlations, but particular for glucose. The computed SEP is 44.7 mg/dl (2.48 mM) after this off-set correction.

#### 4. Conclusion

By extending the work of Hall and Pollard [16], we have found that the combination region contains sufficient spectral information for reliable measurement of all analytes tested except lactate. Sub-millimolar levels of lactate cannot be measured from near-IR spectra collected with a 2.5 mm optical path length. Furthermore, the successful generation of models from spectra corresponding to 242 different serum sample matrices suggests the feasibility of global calibration models for clinical analytes, including glucose. A strong correlation is observed between predicted and actual glucose levels in samples collected nineteen months after samples used to train the glucose calibration model. This correlation indicates that glucose models are based on glucose specific information and that these models are potentially robust over time. Prediction errors for glucose are

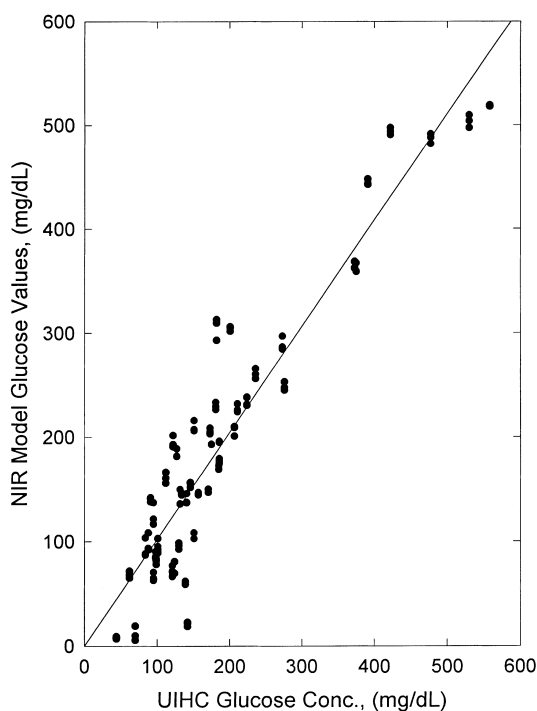


Fig. 9. Prediction correlation plot for blind serum samples after off-set correction. Solid line shows linear regression fit.

not related to total protein levels, lipemic index, or order of analysis.

### Acknowledgements

The contributions and assistance provided by Jason Burmeister and Ronald Shaffer are greatly appreciated. We also thank Professor Ronald Feld, Director of the Clinical Chemistry Laboratory at the University of Iowa Hospitals and Clinics, for collecting all serum samples and holding glucose values for the blind samples. Financial support for this work was provided from the National Institute on Diabetes and Digestive and Kidney Diseases (DK45126).

### References

- [1] M.A. Arnold, New developments and clinical impact of noninvasive monitoring, in: G.J. Kost (Ed.), *Handbook of Clinical Laboratory Automation, Robotics, and Optimization*, Chapter 26, Wiley, New York, 1996.
- [2] H.M. Heise, R. Marbach, T.H. Koschinsky, F.A. Gries, *Appl. Spect.* 48 (1994) 85.
- [3] J. Hall, A. Pollard, *Clin. Chem.* 38 (1992) 1623.
- [4] M.A. Arnold, *Current Opinion Biotechnol.* 7 (1996) 46.
- [5] H.M. Heise, *Horm. Metab. Res.* 28 (1996) 527.
- [6] B. Bauer, T.A. Floyd, *Anal. Chim. Acta* 197 (1987) 295.
- [7] R.M. Gendreau, R.J. Jakobsen, *Biomed. Mater. Res.* 13 (1979) 893.
- [8] R.M. Gendreau, S. Winters, R.I. Leininger, D. Fink, C.R. Hassler, R.J. Jakobsen, *Appl. Spect.* 35 (1981) 353.
- [9] G. Janatsch, J.D. Kruse-Jarres, *Anal. Chem.* 61 (1989) 2016.
- [10] J.D. Kruse-Jarres, G. Janatsch, U. Gless, *Clin. Chem.* 35 (1989) 1845.
- [11] H. Zeller, P. Novak, R. Landgraf, *IJAO* 12 (1989) 129.
- [12] K.L. Ward, D.M. Haaland, M.R. Robinson, R.P. Eaton, *Appl. Spect.* 46 (1992) 959.
- [13] H.M. Heise, R. Marbach, G. Janatsch, J.D. Kruse-Jarres, *Anal. Chem.* 61 (1989) 2009.
- [14] P. Bhandare, Y. Mendelson, R.A. Peura, G. Janatsch, J.D. Kruse-Jarres, R. Marbach, H.M. Heise, *Appl. Spect.* 47 (1993) 1214.
- [15] Y. Mendelson, A.C. Clermont, R.A. Peura, B. Lin, *IEEE Trans. Biomed. Eng.* 37 (1990) 458.
- [16] J. Hall, A. Pollard, *Clin. Biochem.* 26 (1993) 483.
- [17] M.A. Arnold, G.W. Small, *Anal. Chem.* 62 (1990) 1457.
- [18] L.A. Marquardt, M.A. Arnold, G.W. Small, *Anal. Chem.* 65 (1993) 3271.
- [19] G.W. Small, L.A. Marquardt, M.A. Arnold, *Anal. Chem.* 65 (1993) 3279.
- [20] K.H. Hazen, M.A. Arnold, G.W. Small, *Appl. Spect.* 48 (1994) 477.
- [21] H. Chung, M.A. Arnold, M. Rhiel, D.W. Murhammer, *Appl. Biochem. Biotechnol.* 50 (1995) 109.
- [22] H. Chung, M.A. Arnold, M. Rhiel, D.W. Murhammer, *Appl. Spect.* 50 (1996) 270.
- [23] S. Pan, H. Chung, M.A. Arnold, G.W. Small, *Anal. Chem.* 68 (1996) 1124.
- [24] R.E. Shaffer, G.W. Small, M.A. Arnold, *Anal. Chem.* 68 (1996) 2663.
- [25] M.R. Riley, M. Rhiel, X. Zhou, M.A. Arnold, D.W. Murhammer, *Biotechnol. Bioeng.* 55 (1997) 11.
- [26] K.H. Hazen, Ph.D. Dissertation, University of Iowa, 1995.
- [27] J.J. Burmeister, M.A. Arnold, *Anal. Lett.* 28 (1995) 581.
- [28] X. Zhou, H. Chung, M.A. Arnold, M. Rhiel, D.W. Murhammer, Selective measurement of glutamine and asparagine in aqueous media by near-infrared spectroscopy, in: K.R. Rogers, A. Mulchandani, W. Zhou (Eds.), *Biosensor and Chemical Sensor Technology: Process Monitoring and Control*, ACS Symposium Series, vol. 613, Chapter 12, ACS Books, Washington, DC, 1995.
- [29] A. Koller, L.A. Kaplan, Total serum protein, in: A.J. Pesce, L.A. Kaplan (Eds.), *Methods in Clinical Chemistry*, Chapter 147, C.V. Mosby Co., St. Louis, MO, 1987, p. 1134.
- [30] W.L. Clarke, D. Cox, L.A. Gonder-Frederick, W. Carter, S.L. Pohl, *Diabetes Care* 10 (1987) 622.