METHODS

Application of Mid-Infrared Molecular Spectroscopy for Assessment of Biochemical Parameters of Blood Serum

S. A. Khaustova, I. I. Davydov, E. V. Trushkin, M. U. Shkurnikov, R. Mueller*, J. Backhaus* and A. G. Tonevitsky

Translated from *Byulleten' Eksperimental'noi Biologii i Meditsiny*, Vol. 148, No. 12, pp. 707-711, December, 2009
Original article submitted July 9, 2009

The method of mid-infrared molecular spectroscopy allows precise measuring of the concentrations of a large number of biological molecules in a minimal sample volume. Method of projections on latent structures was used for plotting the calibration models. On the basis on mid-infrared spectral data we obtained calibration models for calculation of serum content of various substances: albumin, cholesterol, glucose, total protein, urea, 70 Da heat shock protein, and malonic dialdehyde.

Key Words: *molecular spectroscopy; method of projections on latent structures; serum*

The search for precise, reproducible, and simple methods of simultaneous analysis of concentration of several substances in a small volume of biological material is a priority problem of biological analysis.

Various methods are currently used for measuring the concentration of the test substances in medical biochemistry. These methods have some advantages and drawbacks depending on the type of studied molecules, required accuracy of the analysis, and simplicity. In contrast to these methods, molecular spectroscopy in the mid-infrared spectral range (IR), so-called range of molecular fingerprints, is a sensitive, rapid and easy-to-use physicochemical method enabling analysis of various types of biological samples [10,11]. It is known that the molecule have typical "trace", absorption bands in the medium infrared region, and this makes it possible to identify and characterize each sample and each substance. The method is also used for quantitative analysis of components of complex biological fluids, such as plasma, serum, saliva, urine, and sweat [2-4,6,7,11].

In a series of studies, IR-Fourier spectroscopy was used for measuring the concentration of various substances, *e.g.* glucose, triglycerides, urea, immunoglobulins, lactate *etc.* [1,9,10,13]. These methods require complicated preparation of the analyzed sample [10] and/or multistage modification of the spectrum [8], which can distort the results and reduce reliability of experimental data.

Here we performed preliminary study of the use of IR-Fourier spectroscopy along with other biochemical methods for measuring serum concentration of various substances and for plotting the model for prediction of their serum concentrations.

MATERIALS AND METHODS

Sixteen anonymous blood serum samples were supplied by Laboratory for Testing of Physical Condition, All-Russian Research Institute of Physical Culture and

All-Russian Research Institute of Physical Culture and Sports Education, Moscow, Russia; *Institute for Instrumental Analysis and Bioanalysis, Mannheim University of Applied Sciences, Germany. ***Address for correspondence:*** svetakhaustova@gmail.com. S. A. Khaustova

0007-4888/09/1486-0943

Sports Education (experiment was approved by local ethical committee). Each sample was divided into two aliquots: one aliquot was used for routine biochemical analysis, the other aliquot was analyzed by IR-Fourier spectroscopy. Serum samples were stored at -20°C.

Spectrometry was performed using Tensor 37HTS-XT IR-Fourier spectrometer (Bruker-Optics). Serum samples were diluted 1:4 with distilled water, which allowed to avoid signal saturation in spectral range 1800-1600 cm^{-1}. For optimal spectral detection, 1 μl sample was diluted and homogenized and then placed on the surface of measuring cell of a 384-well silicon plate (Si-Platte HTS-XT, Bruker-Optics) using an automatic doser. For each sample, the assay was done in several repeats. Empty cells (without samples) were used for background spectrum setting. After application of samples, the plate was dried for 12 h in a thermostat in order to remove water traces, after that it was put into the spectrometer. Spectral resolution was 4 cm^{-1}, number of averaging scans was 64 scans per sample, 64 scans per background, the range of wave numbers was 4000-570 cm^{-1} for all spectrometric assays. The experiment was performed three times per day for each sample (9 times during 3 days).

Biochemical analysis was performed using automatic biochemical analyzer HumaStar 300 (Human). IR-Fourier spectrometry was performed using the second aliquot of the analyzed serum. Heat shock protein 70 Da (HSP-70) detection was performed using EIA test system developed on the basis of affinity-purified polyclonal antibodies to HSP-70 [14]. MDA concentration was determined according to previously described protocol [5].

OPUS 5.5 TMS (2005 Bruker Optik GmbH) software installed in IR-Fourier spectrometer (Tensor 37HTS-XT, Bruker) enables automatic spectra processing. Calculation of concentrations of the analyzed substances and model plotting were performed using PLS method of Opus-Quant2 Analysis, which uses reference concentrations and concentrations calculated from the spectrum for calibration model plotting. Based on the dot matrix analysis of the whole spectrum, the method of projections on latent structures (PLS), which uses PLS1 algorithm, allows to establish a better correlation of spectral matrix to known concentrations of the analyzed molecules [15]. Nine spectra were obtained for each serum sample. The assay included samples spectra differing by less than 2% from each other. Position of the sample on the plate was justified by application of the probe on different cells of the plate using automatic dosing apparatus, repeated thrice. Thus, 16×9 spectra were used for model plotting.

Since serum molecules have several specific absorption bands, PLS1 algorithm uses several spectral windows with correlations of specific spectrum changes to values of known concentrations. Therefore, it is crucial to choose a correct set of windows for the analyzed molecule so that the model would possess high predictive capacity.

Normal functioning of the model requires four conditions: a) correct choice of spectral region for the analyzed molecule; b) correlation coefficient R^2 should be as high as possible; c) root mean squared error of cross-validation RMSECV should be as low as possible (RMSECV is a quantitative estimate of the predictive capability of the model, calibration accuracy); d) there must be no more than 10 PLS of regression factors (since molecule concentration depends on the integral effects of various factors, then, the lower is their number, the more precisely the concentration is predicted). R^2 and RMSECV values are automatically kept in memory of the model and are available for the analysis of new samples.

RESULTS

IR-Fourier spectrum is a superposition of spectra of all molecules and compounds present in the sample, the intensity of absorption bands in each individual spectrum is proportional to molecule concentration in the sample (Beer–Lambert–Bouguer law). Serum spectra contain a number of absorption bands typical of the molecules from biological samples (Fig. 1), each band is associated with a family of compounds depending on the presence of specific chemical bonds. Absorption bands in the region between 3020 and 3000 cm^{-1} and between 1732 and 1720 cm^{-1} are specific for cholesterol esters (valence oscillation of methyl, methylene and carbonyl groups). Protein-specific amide 1 (carbonyl group valence oscillations) and amide 2 (N-H group deformation oscillations) peaks are located between 1720 and 1600 cm^{-1} and between 1600 and 1480 cm^{-1}, respectively; absorption regions, specific for saccharides and many other compounds (oscillations of C-O and C-O-C groups) lie between 1300 and 900 cm^{-1} [10].

OPUS-Quant 2 software suggests a number of models, which represent various sets of spectral regions for the detected molecule with different RMSECV values and number of factors. We chose a model with maximal R^2 value, lowest RMSECV value, and one or several spectral windows corresponding to the molecule structure (Fig. 1). For example, HSP-70 calibration model is based on cross-correlation analysis of the second derivative from the obtained serum spectra in region 1591-1478 cm^{-1} and reference HSP-70 concentrations measured by EIA. Selection of this spectral range is dictated by the presence of absorption bands specific for HSP-70 in this region of spectrum

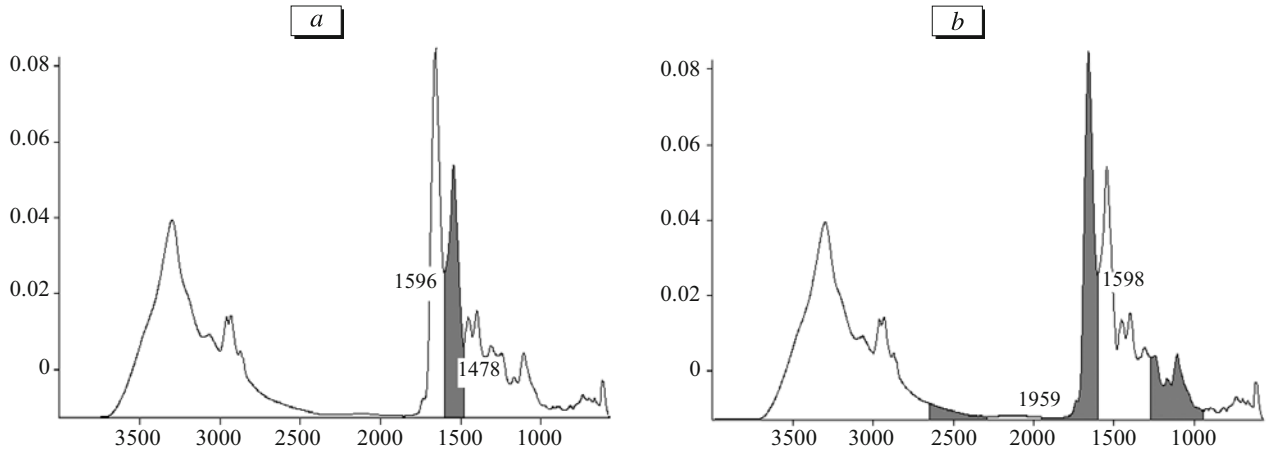

Fig. 1. IR-Fourier spectrum of the serum: characteristic regions for plotting the calibration models for HSP-70 (*a*) and MDA (*b*). Abscissa: wave number, cm^{-1}. Ordinate: absorption.

(deformation oscillations of N-H group with maximum at wave number of 1544 cm^{-1}).

Table 1 shows R^2, RMSECV, range of measured concentrations, and spectral regions for calibration model for each analyzed molecule. Calibration models are specific for each molecule and have individual R^2 and RMSECV values. Calibration model for measuring glucose concentration has high R^2=94 value; glucose concentration determined using this model can differ from the value measured using traditional biochemical method by RMSECV=0.759 mmol/liter (range of concentrations, measured using standard biochemical concentrations method, was 2.6-9.58 mmol/liter). When measuring total protein and albumin concentrations, R^2 were not high: 88 and 84, respectively; however, their concentration can be measured precisely enough: (69.3-82.8)±1.19 g/liter for total protein and (47-53)±0.69 g/liter for albumin.

The concentrations predicted using the model differ from calibration concentrations calculated using standard biochemical methods by less than 4%. Figure 2 presents calibration models for MDA, urea, and HSP-70. The predicted and calibration concentration values are similar and lie along a straight line, which is indicative of good predictive capacity of the model.

TABLE 1. Parameters of Calibration Models

Index	R^2	RMSECV	Concentration range	Spectral region, cm^{-1}
Creatinine	74	8.38 μmol/liter	80-138 μmol/liter	3313-2284; 1256-912
Albumin	84	0.69 g/liter	47-53 g/liter	1602-1478
Amylase	77	38.3 U/liter	59-387 U/liter	1942-1597; 1256-750
Bilirubin	69	1.93 μmol/liter	10.2-22 μmol/liter	3656-3312; 2970-2626; 1598-750
Cholesterol	89	0.242 mmol/liter	3.65-6.27 mmol/liter	3313-2626; 2284-1941
γ-Glutamyltransferase	77	3.04 U/liter	17-39 U/liter	4000-2970; 2284-1941
Glucose	94	0.759 mmol/liter	2.6-9.58 mmol/liter	1178-964
Total protein	88	1.19 g/liter	69.3-82.8 g/liter	4000-3312; 2970-1941; 1256-912
Triglycerides	87	0.114 mmol/liter	0.43-1.54 mmol/liter	4000-1597; 1256-750
Urea	94	0.462 mmol/liter	3.0-11.3 mmol/liter	1833-1478
Testosterone	75	2.29 nmol/liter	4.4-20.5 nmol/liter	3656-2284
Cortisol	89	87.8 nmol/liter	254.1-1100.0 nmol/liter	4000-2300; 1942-1597; 1256-750
Lactate dehydrogenase	79	0.287 U/liter	2.0-4.2 U/liter	2970-2626; 1942-1597; 1256-912
MDA	78	1.08 μmol/liter	12.9-21.3 μmol/liter	2627-2284; 1942-1597; 1256-912
HSP-70	95	1.83 ng/liter	1.8-25.8 ng/liter	1591-1478

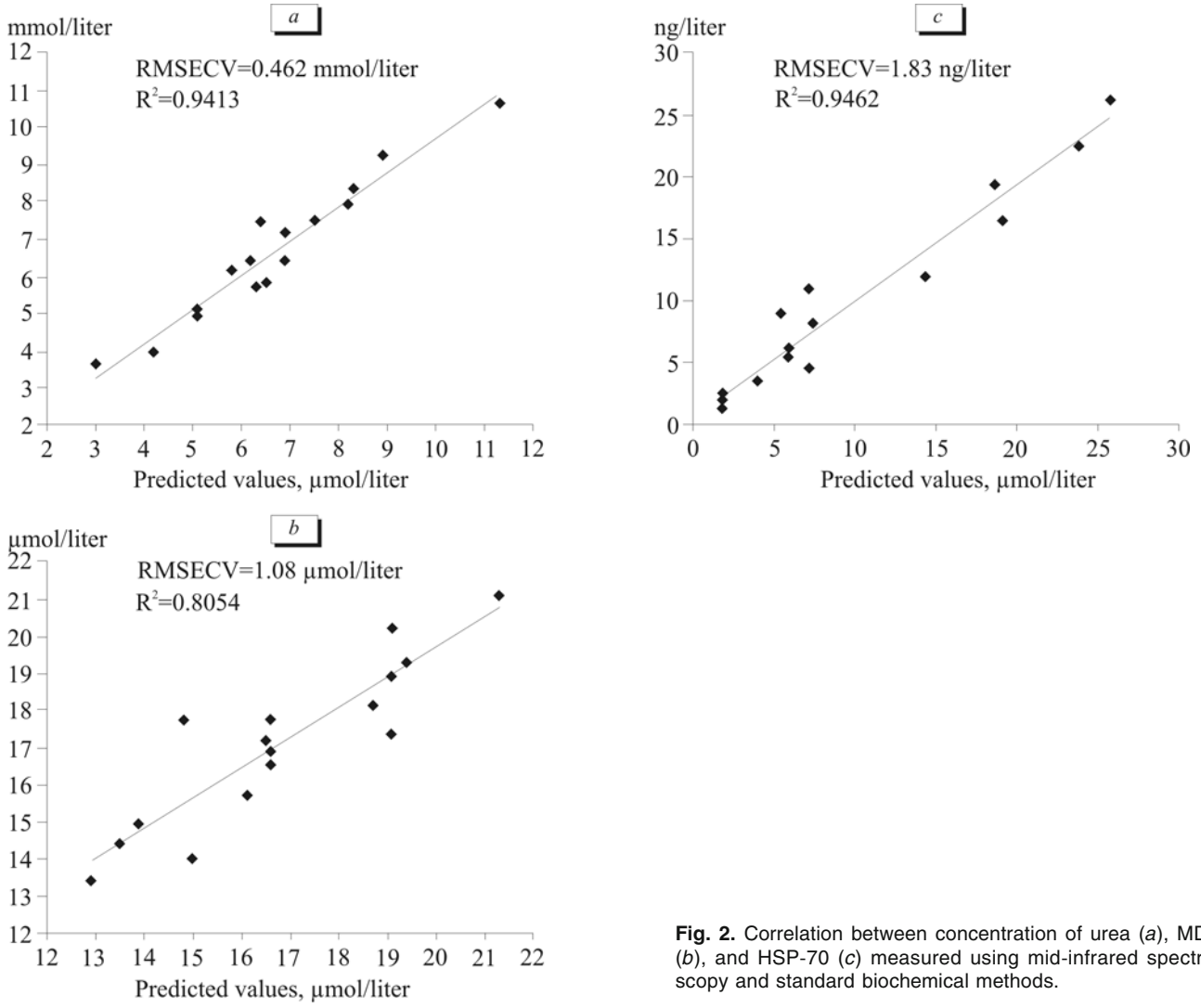

Fig. 2. Correlation between concentration of urea (*a*), MDA (*b*), and HSP-70 (*c*) measured using mid-infrared spectroscopy and standard biochemical methods.

In order to enhance experiment precision, either internal standards are routinely used, or the sample is carefully prepared and mathematical spectrum artifacts are minimized. We preferred IR-Fourier spectroscopy and method of projections on latent structures. On the basis of 144 spectra obtained for 16 serum samples, it was shown that plotting of calibration models with high R^2 and low RMSECV values is possible even for this small sample. This satisfactory result can also be explained by the use of automatic probe application on the plate and maximum similarity of probe processing conditions.

The results of this study make possible a sufficient forward step in measuring the concentration other substances in blood serum. The study takes into account the majority of limitations of techniques used previously and described in published reports; the number of limiting factors hampering measurement of concentration of many substances in one serum probe is reduced as much as possible. First, sample drying under the same conditions for a set of samples was prolonged in order to avoid problems of water absorption in the near-infrared region and approximate preparation conditions as much as possible for the whole set of samples. Second, not only the principal absorption peak, but also absorption in some regions specific for certain molecule was taken into account. Third, the operation of calculation of the absorption peak area for the molecule with subsequent subtraction of the spectrum of pure components from the analyzed serum spectrum was excluded, which excluded the necessity of using reference spectra of pure components [10,11].

On the basis of the obtained data, calibration models were plotted, which make possible determination of concentration of substances in human blood serum using IR molecular spectroscopy (glucose, urea, albumin, cholesterol) in relatively small probe volume (~10 μl). Calibration models for calculation of blood serum content of creatinine, amylase, HSP-70, testosterone, cortisol, and MDA were obtained for the first time.

Relatively low R^2 value can be explained by scarcity of statistical data used for model plotting. We plan to plot calibration model using ample statistical data and evaluate its reliability and reproducibility (model testing with control samples). The obtained spectral data contain information about all substances present in the serum, which allows obtaining information on the majority of substances with sufficient accuracy in one serum aliquot.

The study was supported by Federal Target Program "Investigation and Development in Priority Fields of Scientific and Technological Complex Development of Russia in 2007-2012" (Government contract No. 02.527.11.9024 issued 26 November 2007) and German Federal Ministry of Education and Research (project: DLR RUS No. 08/A13).

REFERENCES

1. D. A. Sakharov, M. U. Shkurnikov, A. V. Stepanov, and A. G. Tonevitsky, *Bull. Exper. Biol.,* **147**, No. 3, 335-340 (2009).
2. K. Ebensen. *Analysis of Multidimetional Data* [in Russian]. Moscow (2005).
3. L. Benezzeddine-Boussaidi, G. Cazorla, and A. M. Melin, *Clin. Chem. Lab. Med.,* **47**, No. 1, 83-90 (2009).
4. J. T. Borden, A. Man, D. A. Scott, and K.-Z. Liu, *J. Mol. Med.*, **81**, No. 12, 788-794 (2003).
5. C. E. Christersson, L. Lindh, and T. Arnebrant, *Eur. J. Oral. Sci.,* **108,** No 5, 418-425 (2000).
6. V. R. Kondepati, H. M. Heise, T. Oszinda, *et al*., *Vibrat. Spectrosc.,* **46**, 150-157 (2009).
7. X.-Y. Li and C. K. Chow, *Lipids*, **29**, No. 1, 73-75 (1994).
8. C. Merle and A. Baillet-Guffroy, *Biochim. Biophys. Acta.*, **1788**, No. 5, 1092-1098 (2009).
9. H. Panayiotou and S. Kokot, *Anal. Chim. Acta*, **392**, 223-235 (1999).
10. C. Petibois, G. Cazorla, A. Cassaigne, and G. Déléris, *Clin. Chem*., **47**, No. 4, 730-738 (2001).
11. C. Petibois, G. Cazorla, and G. Déléris, *Appl. Spectrosc*., **56,** No. 1, 10-16 (2002).
12. C. Petibois, K. Gionnet, M. Gonçalves, *et al*., *Analyst.*, **131**, No. 5, 640-647 (2006).
13. O. Preisner, J. A. Lopes, R. Guiomar, *et al.*, *Anal. Bioanal. Chem*., **387**, No. 5, 1739-1748 (2007).
14. A. Sakudo, H. Kuratsune, Y. H. Kato, and K. Ikuta, *Clin. Chim. Acta.*, **402**, Nos. 1-2, 75-78 (2009).
15. Y. P. Zhou, L. Xu, L.-J. Tang, *et al*., *Anal. Sci.*, **23**, No. 7, 793-798 (2007).