

Analysis of albumin Raman scattering registration efficiency from different volume and shape cuvette

Anastasia A. Lykina*, Dmitry N. Artemyev, and Ivan A. Bratchenko

Samara National Research University, 34 Moskovskoe shosse, Samara 443086, Russia

* e-mail: zuum63@mail.ru

Abstract. The cuvette shape influence on the Raman registration efficiency at small volumes of albumin (0.15-1 ml) using the PLS model was analysed. This paper examines the smallest measurement error RMSE of 3-6 g/l with a correlation coefficient $R^2 = 0.95$ was achieved for pure Raman spectra extraction by baseline correction with asymmetric least squares smoothing method for aluminum spherical type cuvette-reflectors. Its curvature center was matched with the focal spot of the exciting laser. The error of albumin concentration determination equaled about 5% of albumin normal concentration in a blood plasma using a spherical cuvette with a volume of 0.55 ml. A further decrease of the cuvette volume led to a sharp loss of the method prediction accuracy. © 2017 Journal of Biomedical Photonics & Engineering.

Keywords: albumin; raman spectroscopy; PLS.

Paper #3157 received 7 Feb 2017; revised manuscript received 16 Apr 2017; accepted for publication 7 May 2017; published online 15 May 2017. doi: [10.18287/JBPE17.03.020309](https://doi.org/10.18287/JBPE17.03.020309). [Special Issue. Workshop "Biophotonics" of the XIII all-Russian Youth Samara conference-contest on optics and laser physics].

References

1. B. Blomback, and L. A. Hanson, Plasma proteins, John Willey & Sons (1979). ISBN: 978-0-471-99730-6.
2. C. A. Burtis, E. R. Ashwood, and D. E. Bruns, Tietz Textbook of Clinical Chemistry and Molecular Diagnostics, 5th ed., Elsevier Saunders, Philadelphia (2006). ISBN: 978-0-7216-0189-2.
3. F. D. Keren, Protein Electrophoresis in Clinical Diagnosis, Hodder Arnold, London (2003). ISBN: 978-0-340-81213-6.
4. M. Jermyn, J. Desroches, K. Aubertin, K. St-Arnaud, W.-J. Madore, E. De Montigny, M.-C. Guiot, D. Trudel, B. C. Wilson, and K. Petrecca, "A review of Raman spectroscopy advances with an emphasis on clinical translation challenges in oncology," Phys. Med. Biol. 61(23), R370-R400 (2016).
5. N. C. Dingari, G. L. Horowitz, J. W. Kang, R. R. Dasari, and I. Barman, "Raman Spectroscopy Provides a Powerful Diagnostic Tool for Accurate Determination of Albumin Glycation," PLoS ONE 7(2), 1-11(2012).
6. A. J. Berger, T.-W. Koo, I. Itzkan, G. Horowitz, and M. S. Feld, "Multicomponent blood analysis by near-infrared Raman spectroscopy," Appl Opt. 38(13), 2916-2926 (1999).
7. A. M. K. Enejder, T.-W. Koo, J. Oh, M. Hunter, S. Sasic, M. S. Feld, and G. L. Horowitz, "Blood analysis by Raman spectroscopy," Optics letters 27(22), 2004-2006 (2002).
8. A. Bonifacio, S. Dalla Marta, R. Spizzo, S. Cervo, A. Steffan, A. Colombatti, and V. Sergo, "Surface-enhanced Raman spectroscopy of blood plasma and serum using Ag and Au nanoparticles: A systematic study," Analytical and Bioanalytical Chemistry 406(9), 3271-3275 (2015).
9. Z. O. Kralova, A. Orinak, R. Orinakova, L. Skantarova, and J. Radonak, "Enhanced Detection of Human Plasma Proteins on Nanostructured Silver Surfaces," Nanomater. Nanotechnol. 3(17), 1-8 (2014).
10. D. N. Artemyev, V. P. Zakharov, I. L. Davydkin, J. A. Khristoforova, A. A. Lykina, V. N. Konyukhov, and T. P. Kuzmina, "Measurement of human serum albumin concentration using Raman spectroscopy setup," Opt Quant Electron 48(6), 337 (2016).
11. K. Esbensen, Multivariate Data Analysis – In Practice, 4th ed., CAMO, NJ (2000).
12. S. Wold, M. Sjöstroma, and L. Erikssonb, "PLS-regression: a basic tool of chemometrics," Chemometrics and Intelligent Laboratory Systems 58(2), 109-130 (2001).
13. R. Bro, and A. K. Smilde, "Centering and scaling in component analysis," J. Chemom. 17(1), 16-33 (2003).

14. Ch. Ruffin, and R. L. King, "The analysis of hyperspectral data using Savitzky-Golay filtering-theoretical basis," *IEEE International* 2, 756-758 (1999).
15. S.-J. Baek, A. Park, Y.-J. Ahn, J. Choo, "Baseline correction using asymmetrically reweighted penalized least squares smoothing," *The Analyst* 140(1), 250-257 (2015).

1 Introduction

Blood plasma comprises more than a hundred different proteins. Albumin is the most common component of blood plasma (up to 60%), its normal concentration being 35-50 g/l. It acts as a transport protein for some hormones, free fatty acids, bilirubin, supports a colloid-osmotic constancy of blood [1]. The albumin concentration changes indicate the presence of body inflammation or organ pathology.

Physicochemical methods with various reagents are usually used in clinic-based studies to determine the protein fraction concentrations in blood plasma [2]. Protein electrophoresis is one of the most common methods [3], which measures specific proteins in blood. The test separates proteins in blood on the basis of their electrical charges. After the proteins separating, the plate is placed in a solution to stain the protein bands. The staining intensity is related to the protein concentration. Blood samples with a volume of up to 5 ml are required for the protein electrophoresis. The analysis time is 1.5 to 2 hours.

In modern studies, spectral methods are utilised to reduce the protein analysis time. The methods are based on measuring optical properties of substances without additional chemical reagents. Raman spectroscopy (RS) is one of the most sensitive methods [4-6]. The main limitation of RS methods is low scattering cross section, which leads to low Raman scattering intensity. Resonance RS methods, confocal microscopic systems [4-7], surface-enhanced Raman spectroscopy (SERS) using metallic nanoparticles [8, 9] are used to increase the signal-to-noise ratio. However, these technologies are expensive and time-consuming, which impedes their use for routine analysis in medical laboratories.

Usually, the signal-to-noise ratio increase for non-resonant Raman systems [7, 10] is achieved by the signal accumulation (normal law of noise distribution). However, this effect has limitations associated with the signal saturation, for example, in [7] the registration time was 5 minutes. We propose that, the efficiency of laser radiation pumping by means of multireflection and multiple scattering in the focus area of the Raman probe be increased by using cuvette-reflectors with a certain shape. For these purposes, a quantitative analysis of albumin Raman scattering registration efficiency was performed by using cuvette-reflectors with different shape and volume (0.15-1 ml). Since most proteins, including albumin, are natural fluorophores, they have intrinsic fluorescence signal. The albumin limiting concentration also depends on the autofluorescence background removal method to obtain the pure Raman spectrum. We compared the influence of polynomial approximation and baseline correction with asymmetric

least squares smoothing on the prediction accuracy of albumin concentrations.

2 Materials and Methods

2.1 Experimental setup

The optical scheme of the experimental setup for Raman scattering registration is presented in Fig. 1. The InPhotonics RPB785 Raman probe was used to deliver probing laser radiation and collect scattered radiation. It includes elements for scattered radiation filtering. The probing laser radiation (785 nm) was delivered through fiber and a band-pass filter (BPF), which cut off the fluorescent contribution of the optical fiber, and radiation was focused by the L2 lens in the cuvette center (focal length 7.5 mm). The radiation power at the output of the L2 lens was 200 mW. The scattered radiation was collected by the same lens and directed by a dichroic (DM) and deviating (M) mirror to the detection channel through a long-pass filter (LPF) to eliminate Rayleigh component (785 nm) of the scattered radiation. The radiation (fluorescence and Raman scattering) was recorded by QE65Pro Ocean Optics spectrometer with Hamamatsu S7031-1006 detector with a resolution of 3 cm⁻¹. The registered signal was transmitted to the computer for further processing of the Raman spectra.

2.2 Study samples

The medical drug – 10% solution of human albumin (anatomical therapeutic chemical classification system B05AA01) was experimentally tested allowing for simulating a medium close to blood plasma with guaranteed reproducible optical properties. Various concentrations of the model media (5-100 g/l with 5 g/l step) were achieved by mixing albumin and distilled water using microdispenser (accuracy 20 ± 0.2 µl).

2.3 Cuvette

In the study, we used specially designed aluminum cuvette with different volume and three types of shapes: cylindrical with a flat bottom, trapezoidal bottom, and spherical bottom. The material was selected due to the high reflectivity and the absence of an intrinsic Raman spectrum. Fig. 2 shows cuvette with volumes of 0.15 ml, 0.35 ml, 0.55 ml, 0.9 ml and 1 ml.

The depth of the cylindrical cuvettes with a flat bottom was 45 mm and significantly exceeded the diameter (Ø 2 mm, volume 0.35 ml – Fig. 2a and Ø 5 mm, volume 0.9 ml – Fig. 2b) to increase the "light" volume due to the laser radiation reflection from the

side wall surface. The depth of trapezoidal bottom cuvette used as an inverse reflector was selected to commensurate with its “effective” diameter (volume of 1 ml – Fig. 2c). In the spherical type cuvette with volumes of 0.15 ml (diameter of the hole \varnothing 4 mm – Fig. 2d) and 0.55 ml (diameter of the hole \varnothing 8 mm – Fig. 2e) the center of bottom curvature was matched with the focal spot by the internal cavity to fix face plane position of the RS probe. It should be noted that the focal spot is displaced due to multiple scattering in the medium and its size more than 1 mm with an inhomogeneous energy distribution. The spot center is displaced closer to the source due to absorption. This geometry increases the light field energy concentration in the focal area by reflecting from the cuvette bottom and collects scattered forward radiation.

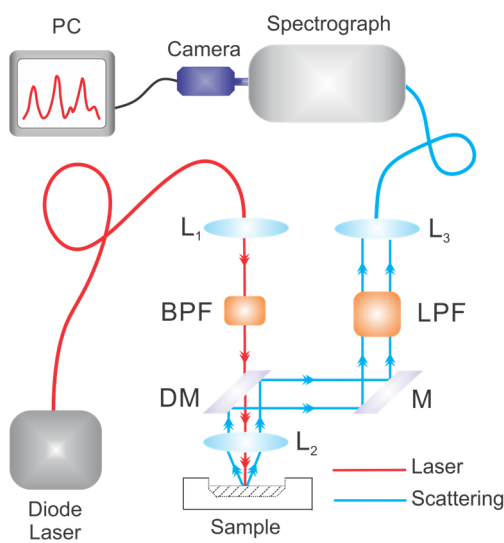


Fig.1 Optical scheme of the experimental setup.

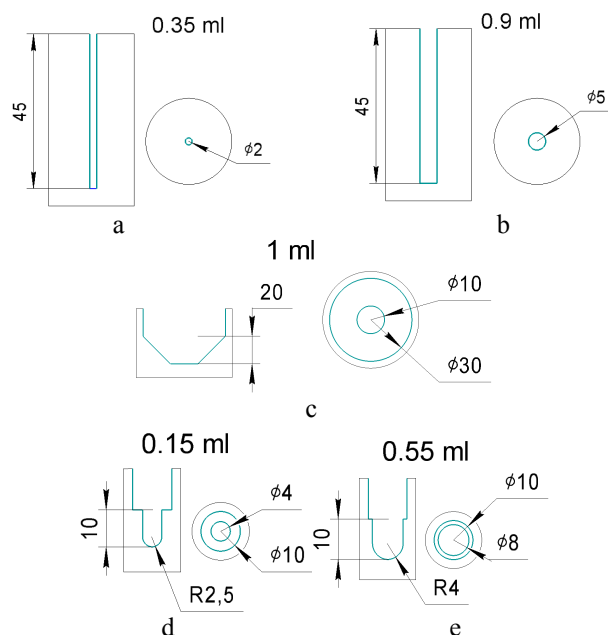


Fig. 2 Cuvettes with volumes of 0.35 ml (a), 0.9 ml (b), 1 ml (c), 0.15 ml (d), 0.55 ml (e).

2.4 Data processing methods

Projection on latent structures (PLS) method was used for the experimental data analysis [11, 12], which interprets the results based on a smaller number of bilinear components. Its goal is to predict a set of dependent variables (vector Y – albumin concentrations) from a set of independent variables or predictors (matrix X – albumin Raman spectra intensities for different wavelengths). This prediction is achieved by extracting from predictors a set of orthogonal factors called latent variables (LV) which have best predictive power.

Model performances were characterised by the root mean square error (RMSE) of calibration, prediction or cross-validation, as well as the coefficient of determination R^2 .

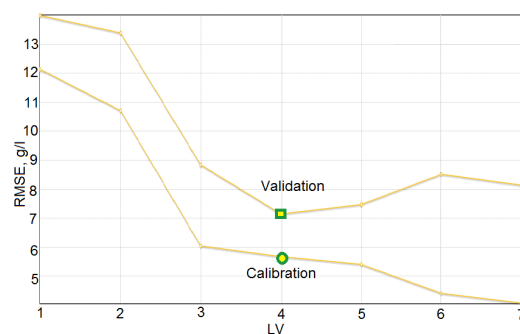


Fig. 3 An example plot of the root mean square error of calibration ($RMSE_C$) versus root mean square error of cross validation ($RMSE_{CV}$) as a function of the number of latent variables. The optimum number of latent variables corresponds to the minimum of $RMSE_{CV}$.

Fig. 3 illustrates typical RMSE dependence with latent variables increase in the analytical model. It is observed that a certain number of LV (in this example $LV = 4$) corresponds to the minimum value of RMSE of cross validation ($RMSE_{CV}$). The number of LV was determined and used to evaluate the model for each cuvette.

The registered signal includes the autofluorescence and Raman spectra, so a raw signal preprocessing was performed to background (autofluorescence) removal. Two methods of the Raman spectrum extraction were compared: polynomial approximation method and baseline correction with asymmetric least squares based on procedures implemented in the TPTcloud cloud service (<https://tptcloud.com/>). The baseline correction approach includes mean centering and data normalization using standard normal variate (snv) method [12, 13]. Savitzky-Golay (sav gol) filter was used to noise reduction [14], and background elimination was realised by baseline correction with asymmetric least squares (baseline als) [15]. Typical results of preprocessing methods application are shown in Fig. 4.

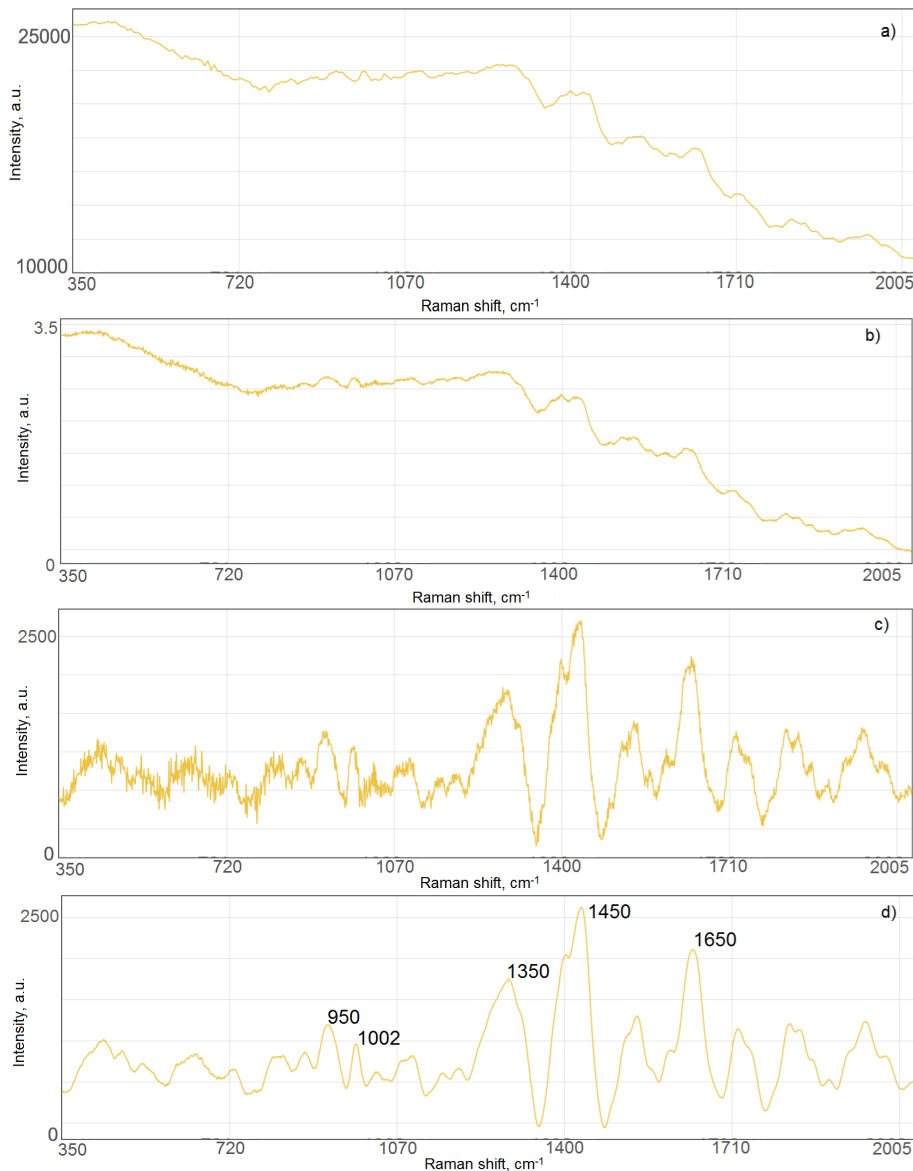


Fig. 4 Preprocessing methods results for albumin RS signals: a) raw signal, b) standard normal variate (snv), c) baseline correction with asymmetric least squares (baseline als), d) Savitzky-Golay filter (sav gol).

The polynomial method utilisation for the Raman spectra fitting is shown in Fig. 5 which includes averaging and smoothing (Fig. 5b), approximation of the averaged signal by a polynomial of the 10th degree (Fig. 5c) to obtain pure Raman spectrum (Fig. 5d). Then, the pure Raman spectra were also normalised (snv) and centered (mean center).

Validation method was applied to estimate the model performance on “independent” data. KFold cross-validation method was used [11]. The KFold splits the data onto K blocks (10 in our model). Further on, the blocks were excluded (one at a time) from the training data and used for validation of the model built on the remaining data. Validation statistics for all individual segments were then averaged.

3 Results and Discussion

For each cuvette scattering spectra were five times registered with an albumin concentration ranging from 5 to 100 g/l (step of 5 g/l). Each raw spectrum was preprocessed by two methods (polynomial and baseline als), afterwards PLS model was built for spectral data of cuvette. The main peak positions of albumin Raman spectra (Figs 4, 5) are clearly observed at wavenumbers: 950 cm^{-1} (stretching mode of CC bond), 1002 cm^{-1} (phenylalanine), 1350 cm^{-1} (deformation mode of CH bond), 1450 cm^{-1} (deformation mode of CH_2 bond) and 1650 cm^{-1} (Amide I) [5].

Predictive performances of the PLS models show a strong correlation between the concentrations predicted by Raman spectroscopy and target albumin concentrations. The PLS model performances of 0.9 ml cylindrical cuvette were compared for the baseline als

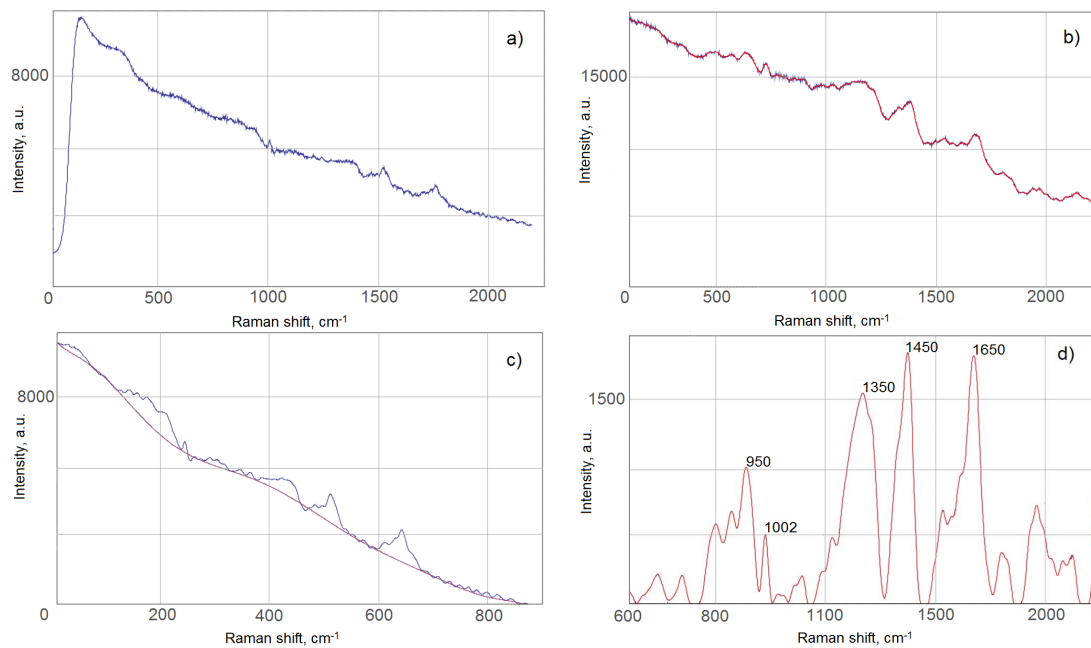


Fig. 5 Raman spectra fitting by polynomial approximation method: a) raw data, b) averaging and smoothing, c) polynomial approximation, d) pure Raman spectrum.

technique (Fig. 6) and for the polynomial technique (Fig. 7). The PLS calibration models show the minimum RMSE with $LV = 4$ for the baseline als method and with $LV = 5$ for polynomial approximation method.

The correlation coefficient (R^2) values above 0.9 were obtained for both processed methods, which indicate the results consistency. However, the baseline als method produces a slightly smaller data spread $\{LV = 4, RMSE = 5.7, R^2 = 0.96, RMSE_{cv} = 7, R^2_{cv} = 0.94\}$ compared to the polynomial method $\{LV = 5, RMSE = 6.5, R^2 = 0.95, RMSE_{cv} = 8.5, R^2_{cv} = 0.92\}$. Here the index “CV” indicates the data obtained by cross-validation. The results of the PLS albumin concentration prediction for all cuvette types are summarised in Table 1.

Comparison of the results shows that the baseline als method for pure Raman spectra extraction has an undoubted advantage over all types of cuvettes, providing smaller values of RMSE.

It is clearly seen that the cylindrical cuvette with extra depth ($\varnothing 5$ mm, volume of 0.9 ml) has similar accuracy characteristics as a shallow wide cuvette with a trapezoidal bottom of the same volume (1 ml). This fact indicates that reflection contribution of forward scattered radiation from the bottom and illumination increase due to reflection from the side surface has the same magnitude order of the measurement error reduction. This effect is particularly significant to the cuvette with the spherical bottom with respect to the curvature center matching with focal spot position of the Raman probe. In this case, it is possible to obtain the same measuring errors of albumin concentration for cuvettes with half volume (0.55 ml): $\{RMSE_{cv} = 6.3, R^2_{cv} = 0.95\}$ compared to cylindrical cuvette with a volume of 0.9 ml $\{RMSE_{cv} = 7, R^2_{cv} = 0.94\}$.

Apparently, volume of 0.5 ml, is critical because a sharp decrease of the correlation coefficient R^2 for cross-validation data is observed for both cylindrical and spherical cuvette as well as for cuvettes with a bottom; R^2 starting from this volume value becomes less than 0.9. Nevertheless, it can be noted that these trends are of different character. For example, the volume increase of the cylindrical cuvette by a factor of 2.6 leads to accuracy increase in albumin concentration determination by 58%, and for spherical type cuvettes by 102% with a volume increase by 3.7 times.

It should be noted that for the spherical type cuvette a small displacement (1 mm) of the bottom curvature center position to the source relative to focal spot center of the Raman probe decreases the determination concentration error.

It could be concerned with the biofluid refraction and the irregular distribution of a radiation energy, associated with absorption of the laser radiation.

It can be argued that the use of the spherical shape cuvette with a volume of ~ 0.5 ml makes it possible to increase the determination accuracy of albumin concentration to 3-6 g/l corresponding to 5-10% of a normal albumin concentration in a blood plasma. This value is comparable to the error of more complicated methods of Raman microscopy and equals about 2 g/l, but this technique requires the use of specialised optics [5].

4 Conclusion

The study shows that for small volume cuvette-reflectors their shape could significantly influence registration efficiency of the albumin Raman signals. In particular, the implementation of the cuvette with a spherical bottom leads to an almost twofold accuracy

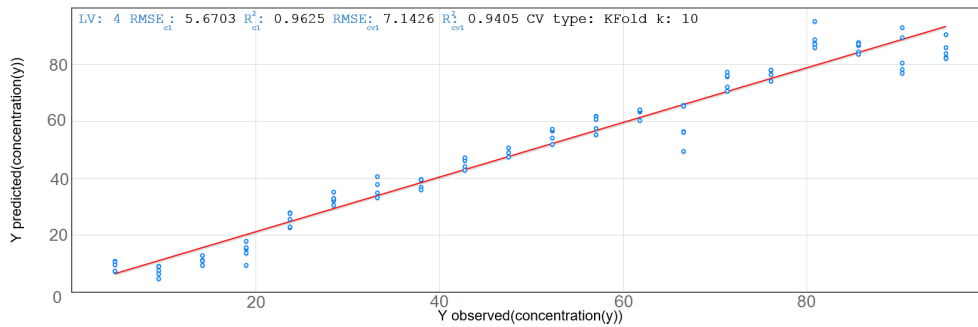


Fig. 6 Prediction results obtained using PLS regression on albumin spectra processed by baseline als, sav gol, snv and mean center for the 0.9 ml cuvette. (LV = 4, RMSE = 5.7, R² = 0.96, RMSE_{cv} = 7 and R²_{cv} = 0.94).

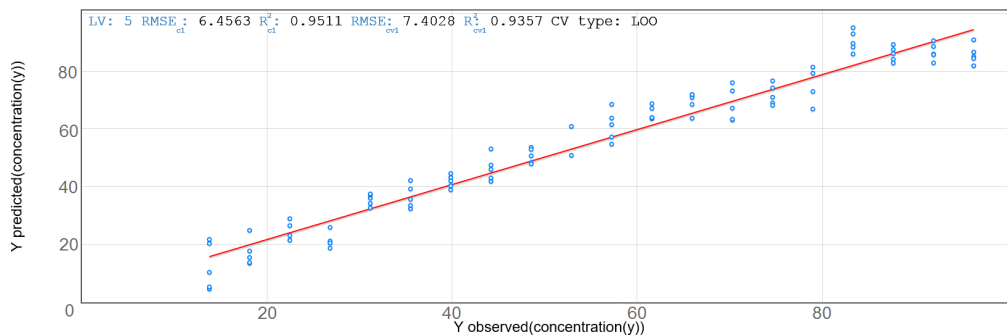


Fig. 7 Prediction results obtained using PLS regression on albumin spectra processed by polynomial approximation, snv and mean center for the 0.9 ml cuvette. (LV = 5, RMSE = 6.5, R² = 0.95, RMSE_{cv} = 8.5 and R²_{cv} = 0.92).

Table 1 The determination accuracy of albumin concentration by the PLS regression methods for cuvette of different shapes and volumes (the highest determination accuracy for the cross-validation data set is in bold type).

Cuvette type	Cylindrical		Trapezoidal	Spherical	
Cuvette volume	0.35 ml	0.9 ml	1 ml	0.15 ml	0.55 ml
Baseline als method					
Calibration data	RMSE = 6.8 R ² = 0.94 (LV = 5)	RMSE = 5.7 R ² = 0.96 (LV = 4)	RMSE = 4.7 R ² = 0.97 (LV = 3)	RMSE = 5.9 R ² = 0.96 (LV = 5)	RMSE = 2.6 R ² = 0.99 (LV = 5)
Cross-validation data	RMSE _{cv} = 10.6 R ² _{cv} = 0.87	RMSE_{cv} = 7 R²_{cv} = 0.94	RMSE _{cv} = 6.1 R ² _{cv} = 0.95	RMSE _{cv} = 14 R ² _{cv} = 0.75	RMSE_{cv} = 6.3 R²_{cv} = 0.95
Polynomial method					
Calibration data	RMSE = 9.5, R ² = 0.89 (LV = 5)	RMSE = 6.5 R ² = 0.95 (LV = 5)	RMSE = 5.2 R ² = 0.96 (LV = 3)	RMSE = 8 R ² = 0.92 (LV = 4)	RMSE = 7.5 R ² = 0.94 (LV = 4)
Cross-validation data	RMSE _{cv} = 11.6 R ² _{cv} = 0.84	RMSE _{cv} = 8.5 R ² _c = 0.92	RMSE _{cv} = 6.7 R ² _{cv} = 0.95	RMSE _{cv} = 20 R ² _c = 0.5	RMSE _{cv} = 11 R ² _{cv} = 0.86

increase due to collection of the scattered forward Raman radiation. It was stated that RMSE of the PLS model is less than 1-5 g/l for all types of cuvettes using baseline correction with asymmetric least squares method to obtain the pure Raman spectra in comparison with the polynomial method. The smallest error (3-6 g/l) was achieved for the spherical type cuvette with a volume of 0.55 ml by matching the center of curvature

with the Raman probe focus. This accuracy allows for reliable determination of abnormal albumin concentrations in blood plasma. The volume decrease less than 0.5 ml (0.35 ml and 0.15 ml) for any shape cuvettes leads to a sharp drop of correlation coefficient (R² << 0.9), which indicates failure to obtain satisfactory results without utilization of SERS-substrates, resonant or confocal microscopic systems.

5 Conclusion

The study shows that for small volume cuvette-reflectors their shape could significantly influence registration efficiency of the albumin Raman signals. In particular, the implementation of the cuvette with a spherical bottom leads to an almost twofold accuracy increase due to collection of the scattered forward Raman radiation. It was stated that RMSE of the PLS model is less than 1-5 g/l for all types of cuvettes using baseline correction with asymmetric least squares method to obtain the pure Raman spectra in comparison with the polynomial method. The smallest error (3-6 g/l) was achieved for the spherical type cuvette with a

volume of 0.55 ml by matching the center of curvature with the Raman probe focus. This accuracy allows for reliable determination of abnormal albumin concentrations in blood plasma. The volume decrease less than 0.5 ml (0.35 ml and 0.15 ml) for any shape cuvettes leads to a sharp drop of correlation coefficient ($R^2 \ll 0.9$), which indicates failure to obtain satisfactory results without utilization of SERS-substrates, resonant or confocal microscopic systems.

Acknowledgments

This research was supported by the Ministry of Education and Science of the Russian Federation.