

Raman spectra analysis of human blood protein fractions using the projection on latent structures method

A.A. Lykina¹, D.N. Artemyev¹, I.A. Bratchenko¹, Yu.A. Khristoforova¹, O.O. Myakinin¹,
T.P. Kuzmina², I.L. Davydkin², V.P. Zakharov¹

¹Samara National Research University, 34 Moskovskoe Shosse, 443086, Samara, Russia

²Samara State Medical University, 89 St. Chapayevskaya, 443099, Samara, Russia

Abstract

This work is devoted to the study of human blood protein fractions by Raman spectroscopy. Whole blood and blood plasma were used as the tested samples. For the pure Raman spectra analysis the autofluorescence background was subtracted by using of two mathematical approaches: polynomial approximation and baseline correction with asymmetric least squares. The study allowed for revealing the differences between the spectral features of blood plasma and whole blood plasma, which are changes in the relative Raman intensities of plasma and whole blood and appearance Raman bands 670 cm^{-1} , 750 cm^{-1} , 1120 cm^{-1} and 1550 cm^{-1} correspond to hemoglobin bonds in whole blood. The spectral features were used for total protein concentration measurement of plasma and whole blood. PLS regression method was utilized for spectral data analysis with different protein concentrations. The VIP-scores make it possible to determine the most informative spectral bands: 1002 cm^{-1} , 1227 cm^{-1} , 1400 cm^{-1} , 1630 cm^{-1} for proteins analysis.

Keywords: whole blood; blood plasma; Raman spectroscopy; Projection on Latent Structures

1. Introduction

Proteins are high molecular weight polypeptides consisting of amino acids and are part of all human body fluids. Changes in the blood protein amount, as well as the certain fractions quantity, allows for drawing a conclusion about the human body state and pathology presence, also it helps to define the treatment efficiency [1].

Various spectral methods allow for obtaining an individual spectral “fingerprint” of the tested sample chemical compounds and these techniques are used for qualitative evaluation of blood protein. Raman spectroscopy (RS) is one of the most sensitive optical methods [2-6], as this approach has been used for various blood proteins analysis [3, 4]. The aim of this study was to compare Raman spectra of plasma and whole blood (mixture of plasma and formed elements, such as erythrocytes) [6]. The application of two mathematical approaches (polynomial approximation and baseline with asymmetric least squares) was demonstrated for Raman signal separation from the autofluorescence background. Regression model of the Projection on Latent Structures (PLS) method was constructed for the determination of total protein concentration by the analysis of plasma and whole blood Raman spectra. Variable importance in projection-scores allow for determining the most informative spectral bands.

2. Material and methods

2.1. Experimental setup

Raman spectra were collected by setup including thermally stabilized diode laser LML-785.0RB-04 (785 nm, 200 mW), commercial Raman probes (Inphotonics RPB785), and spectrograph Shamrock SR-500i-D1-R with deeply cooled digital camera Andor iDus DU416A-LDC-DD (air-cooled up to $-70\text{ }^{\circ}\text{C}$). Detailed information about the experimental setup presented in paper [7].

All spectra were recorded in 780-950 nm spectral range, the exposure time was 60 seconds. The recording of three spectra for each studied sample was performed sequentially. The total time of Raman spectra registration was 3 minutes.

2.2. Tested samples preparation and spectra registration

The standardized collection of whole blood samples from patients with pathological blood disease was performed. The whole blood samples were obtained from the biochemical laboratory of the Samara State Medical University. Blood plasma was produced by sedimentation of whole blood in a test-tube at $+2 + 4^{\circ}\text{C}$ up to the complete drop-out of the formed elements to the bottom of the tube. After the blood plasma sample has been studied, it was mixed with the formed elements for the subsequent analysis of the whole blood. Altogether we performed our study for 45 samples. The tested samples were placed in the aluminum cuvette with a volume of 0.9 ml. Choice of the cuvette geometry was made based on our previous study results

[7]. The cuvette had a cylindrical shape with a flat bottom, the cuvette depth was 45 mm, the hole diameter was \varnothing 5 mm. The chosen cuvette geometry provides the increase of “light” volume due to the laser radiation reflection from the side walls surface.

2.3. Data processing methods

PLS method was used for the experimental data analysis [8], as this method interprets the results based on a smaller number of bilinear components.

The registered signal includes the autofluorescence and Raman components, so a raw spectrum preprocessing was performed for the autofluorescence background removal. Two methods of the Raman spectrum extraction were utilized: polynomial approximation method and baseline correction with asymmetric least squares based on procedures implemented in the TPTcloud (<https://tptcloud.com/>) cloud service. Baseline correction for background component removal was performed using the method of asymmetric least squares (baseline als) [9]. Another approach utilized in this study for Raman spectra separation from the autofluorescence signal is the polynomial approximation [10].

Spectral informative bands during the regression model construction were determined by the analysis of the variable importance in projection (VIP) [11]. The higher the VIP-score of an individual variable corresponds to the more significant values in the constructed model. Variables with a low VIP-score are less important, and may be regarded as candidates for exclusion from the model.

3. Results and discussion

3.1. Raw spectra processing

The raw spectra of plasma and whole blood were registered for the study of protein fractions. The common bands of protein fractions were obtained on the basis of two mathematical approaches: polynomial approximation and method of asymmetric least squares. Analysis of the plasma and whole blood spectra allows for detection of the differences in the proteins component composition. Fig. 1 demonstrates Raman bands of blood plasma 820 cm^{-1} (vibrations of tyrosine), 950 cm^{-1} (deformation vibrations of CH group), 1002 cm^{-1} and 1080 cm^{-1} (vibrations of phenylalanine), 1160 cm^{-1} (deformation vibrations of CC group), 1250 cm^{-1} (α -helix in Amide III), 1330 cm^{-1} (vibrations of tryptophan), 1450 cm^{-1} (deformation vibrations of $-\text{CH}_2$ group), 1650 cm^{-1} (β -helix in Amide I) [12]. Fig. 2 shows Raman spectra of whole blood. The bands are similar to Raman bands of plasma excluding 570 cm^{-1} (deformation vibrations of FeO_2 group), 670 cm^{-1} and 750 cm^{-1} (vibrations of pyrrole), 1120 cm^{-1} (deformation vibrations of C-N group), 1227 cm^{-1} (deformation vibrations of CH group), 1550 cm^{-1} (vibrations of phenylalanine) [13]. Each processed spectrum was a subject to the multivariate analysis for the construction of regression model. The spectra of plasma and whole blood processed by two approaches (baseline als and polynomial approximation) and normalized using the standard deviation are shown in Fig. 1 (b) and Fig. 2 (b).

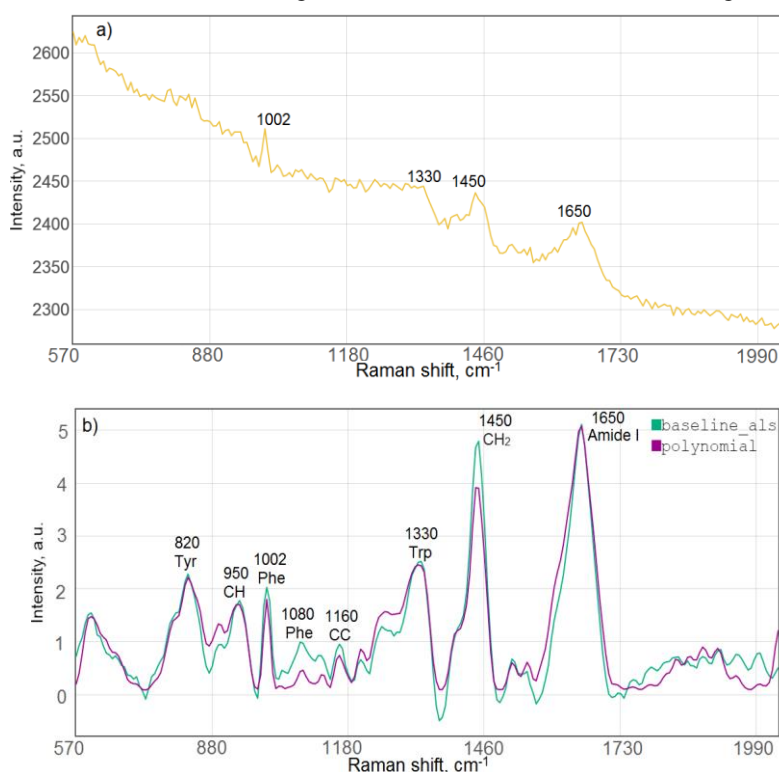


Fig. 1. Raman spectra of blood plasma processed by methods of baseline als, polynomial approximation (Phe- phenylalanine, Trp- tryptophan, Tyr-tyrosine)
a) raw spectrum b) pure Raman spectrum.

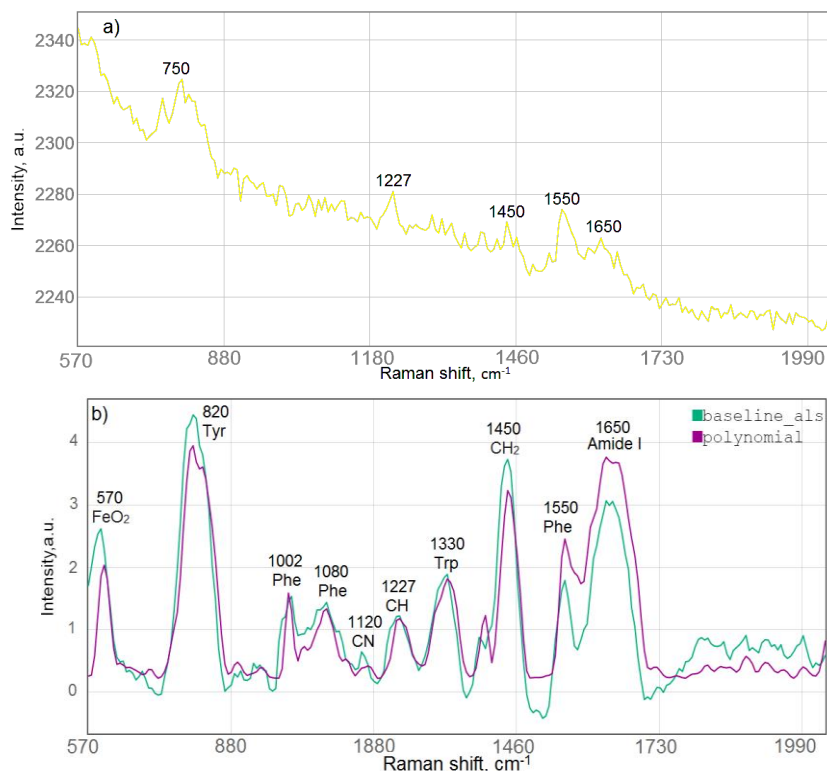


Fig. 2. Raman spectra of whole blood processed by methods of baseline als, polynomial approximation (Phe- phenylalanine, Trp- tryptophan, Tyr-tyrosin)
a) raw spectrum b) pure Raman spectrum.

As shown in Fig.1 utilization of baseline als and polynomial approximation provides the possibility to observe Raman bands corresponding to the contribution of certain molecular vibrations. Analysis of Fig. 1 (b) demonstrate that the shape and intensity of Raman peaks in the blood plasma spectrum processed by baseline als and polynomial approximation are coincide in spectral ranges: 980- 1100 cm^{-1} , 1150-1190 cm^{-1} , 1640-1680 cm^{-1} . For blood plasma samples the intensity of 1080 cm^{-1} , 1160 cm^{-1} and 1450 cm^{-1} Raman bands is 15-20% higher for processing by baseline als algorithm unlike using the polynomial approximation. Analysis of 1700-2000 cm^{-1} spectral region was not performed, since shape of spectra in this region is mostly associated with contribution from the optical filtering module.

Analysis of Fig. 2 helps to conclude, that raw spectra of whole blood processed by two methods Raman bands become more informative due to the elimination of autofluorescence and the appearance of characteristic bands that are hardly recognizable on a raw spectra. Positions of whole blood Raman bands coincide on the entire spectral range. Herewith, the maximum intensity difference on the 1650 cm^{-1} band does not exceed 25%.

3.2. Raman spectra of blood plasma and whole blood

To compare the Raman bands of blood plasma and whole blood a data normalization using standard normal variate (snv) method was performed. Fig. 3 shows the normalized averaged spectra of blood plasma and whole blood for all 45 tested samples.

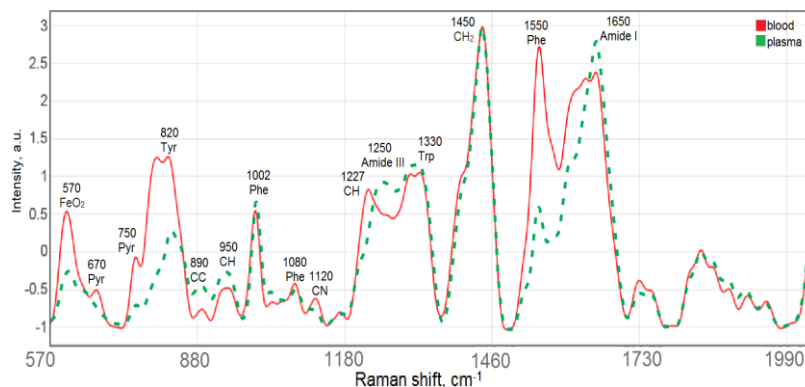


Fig. 3. Normalized averaged pure Raman spectra of blood plasma and whole blood processed by polynomial approximation method (Pyr- pyrrole, Phe- phenylalanine, Trp- tryptophan, Tyr-tyrosin).

Fig. 3 demonstrates that the Raman peaks intensities of blood plasma and whole blood coincide at 1002 cm^{-1} and 1450 cm^{-1} bands. The common band on 1002 cm^{-1} is phenylalanine, which corresponds to a protein amino acid, the precursor of all nutrients [14]. The Raman peak on 1450 cm^{-1} (deformation vibrations of $-\text{CH}_2$ group) is present in both spectra of blood

plasma and whole blood. On the spectral ranges of $570\text{--}800\text{ cm}^{-1}$ and $1460\text{--}1600\text{ cm}^{-1}$, figure shows that intensity of the Raman spectra of whole blood is higher than the intensities for blood plasma. This fact is caused by hemoglobin presence in the whole blood [14]. This is an iron protein, present in erythrocytes [15]. Hemoglobin common Raman peaks are 570 cm^{-1} , 820 cm^{-1} , 670 cm^{-1} , 750 cm^{-1} , 1227 cm^{-1} and 1550 cm^{-1} . The peak on 570 cm^{-1} corresponds to the deformation vibrations of FeO_2 group of hemoglobin. Pyrrole, one of the hemoglobin components have strong Raman peaks at 670 cm^{-1} and 750 cm^{-1} bands. The 1227 cm^{-1} band corresponds to deformation vibrations of CH group of hemoglobin [16]. The Raman bands on 820 cm^{-1} and 1550 cm^{-1} correspond to the vibration of tyrosine and phenylalanine, which contribute to the composition of the protein components of blood plasma and whole blood. Since these amino acids are parts of hemoglobin, the intensity of whole blood tyrosine and phenylalanine Raman bands is 70-75% above than their Raman intensities in the blood plasma.

3.3. PLS analysis of Raman spectra of blood plasma and whole blood

The VIP-scores of Raman spectra matrix of the plasma and whole blood samples for the constructed regression model of the total protein concentration prediction are shown in Fig. 4.

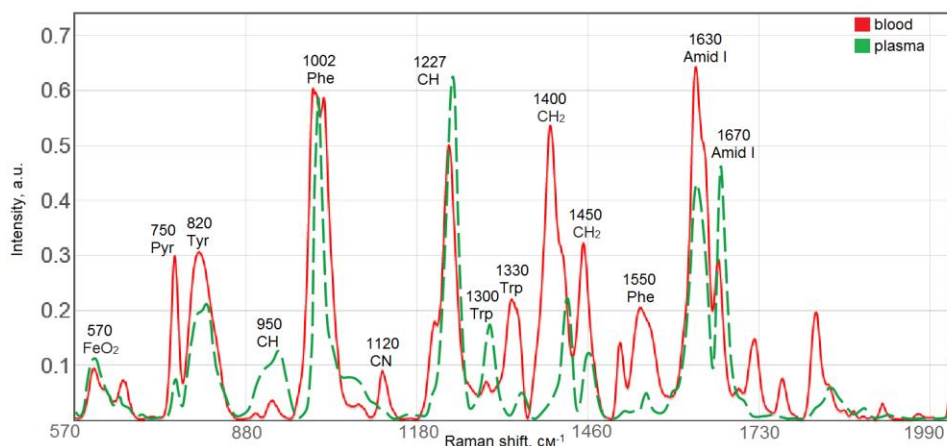


Fig. 4. VIP-scores for PLS multivariate statistical model (Pyr- pyrrole, Phe- phenylalanine, Trp- tryptophan, Tyr-tyrosin).

Fig. 4 demonstrates that the most Raman peaks of plasma and whole blood spectra are coincide on the full spectral range. The most informative spectral bands for whole blood are $570\text{--}700\text{ cm}^{-1}$, 1120 cm^{-1} , 1550 cm^{-1} ; and these peaks are not observed in the blood plasma Raman spectra. These bands are associated with hemoglobin groups. The Raman spectra of plasma and whole blood include multiple peaks in $970\text{--}1040\text{ cm}^{-1}$, $1370\text{--}1500\text{ cm}^{-1}$ and $1580\text{--}1710\text{ cm}^{-1}$ bands, and these peaks are mixed in single peaks, observed in registered spectra: 1002 cm^{-1} , 1450 cm^{-1} and 1650 cm^{-1} . In our study the VIP distribution allows for evaluation of the Raman spectra. It doubles the spectral band, herewith increasing the Raman peaks informativeness. Analysis of obtained results makes it possible to draw a conclusion that the peaks of VIP distribution for the constructed regression model of the total protein concentration predict coincide with Raman peaks shown in Fig.3. Herewith differences are observed in the intensity amplitude of the spectral bands. As shown in Fig. 4 the largest values of VIP-scores corresponds to peaks on spectral bands of 1002 cm^{-1} , 1227 cm^{-1} , 1440 cm^{-1} and 1630 cm^{-1} . The chosen bands correspond to albumin and globulin [17, 18], whose concentration predominates in plasma and whole blood.

4. Conclusion

The current study demonstrates analysis of the plasma and whole blood Raman spectra obtained by two mathematical approaches: polynomial approximation and asymmetric least squares. The maximum differences in the Raman bands intensities did not exceed 20-25% for both approaches.

The carried out research allowed for differences detection between blood plasma and whole blood Raman spectra. The main differences in the spectral characteristics of the tested samples are observed in 670 cm^{-1} , 750 cm^{-1} , 1120 cm^{-1} and 1550 cm^{-1} bands. These bands are associated with hemoglobin bonds, such as pyrrole and FeO_2 vibrations.

The VIP-scores calculation makes it possible to define the most informative spectral bands for total proteins analysis 1002 cm^{-1} , 1227 cm^{-1} , 1400 cm^{-1} , 1630 cm^{-1} corresponding to albumin and globulin fractions.

Acknowledgments

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

References

- [1] Hanlon EB, Manoharan R, Koo TW, Motz JT, Fitzmaurice M, Kramer JR, Itzkan I, Dasar RR, Feld MS. Prospects for in vivo Raman spectroscopy. Luxembourg: Phys. Med. Biol. 2000; 45(2): 59.
- [2] Premasiri WR, Lee JC, Ziegler LD. Surface-Enhanced Raman Scattering of Whole Human Blood, Blood Plasma, and Red Blood Cells: Cellular Processes and Bioanalytical Sensing. Luxembourg: J. Phys. Chem. B. 2012; 116(31): 9376.

- [3] Artemyev DN, Bratchenko IA, Khristoforova JA, Lykina AA, Myakinin OO, Kuzmina TP, Zakharov VP, Davydkin I.L. Blood proteins analysis by Raman spectroscopy method. *Izbrannyye Trudy. – Luxembourg: Proceedings of SPIE-The International Society for Optical Engineering* 2016; 98887:1Y–1.
- [4] Annika MK, Tae-Woong E, Oh J, Hunter M. Blood analysis by Raman spectroscopy. *United States: Optics letters* 2004; 27: 2004.
- [5] Dingari NC, Horowitz GL, Kang JW, Dasari RR, Barman I. Raman Spectroscopy Provides a Powerful Diagnostic Tool for Accurate Determination of Albumin Glycation. *Francisco: PLoS ONE* 2012; 7: 2.
- [6] Castiglioni C, Tommasini M, Zerbi G. Raman spectroscopy of polyconjugated molecules and materials: confinement effect in one and two dimensions. *United States: Philos. Trans. A Math. Phys Eng. Sci*, 2004; 1824: 2469.
- [7] Lykina AA, Artemyev DN, Bratchenko IA. Analysis of albumin Raman scattering registration efficiency from different volume and shape cuvette. *United States: JBPE* 2017; 2: 3157.
- [8] Esbensen KH. *Multivariate Data Analysis. New Jersey: In Practice* 4-th Ed. 2000.
- [9] Eilers PHC, Boelens HFM. Baseline Correction with Asymmetric Least Squares Smoothing. *United States: Leiden University Medical Centre* 2005.
- [10] Zeng H, Lui H, McLean DI. Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy. *United States: Applied spectroscopy* 2007; 61: 1225.
- [11] Farrés M, Platikanov S, Tsakovski S, Tauler R. Comparison of the variable importance in projection (VIP) and of the selectivity ratio (SR) methods for variable selection and interpretation. *United States: Journal of Chemometrics* 2015; 10: 528.
- [12] Regula A, Majzner K, Marzes KM, Kaczor A, Pilarczyk M, Baranska M. Raman spectroscopy of proteins: a review. *United States: J. Raman Spectroscopy* 2017; 44: 1061.
- [13] Atkins CG, Buckley K, Blades MW, Turner RFB. *Raman Spectroscopy of Blood and Blood Components. United States: Applied spectroscopy* 2017; 5: 767.
- [14] Gelder JDe, Gussem KDe, Vandenabeele P, Moens L. Reference database of Raman spectra of biological molecules. *United States: J. Raman Spectroscopy* 2007; 9: 1133.
- [15] Das TK, Counture M, Ouellet Y, Guertin M, Rousseau DL. Simultaneous observation of the O—O and Fe—O₂ stretching modes in oxyhemoglobins. *United States: PNAS* 2011; 5: 479.
- [16] Casella M, Lucotti A, Tommasini M, Zerbi G. Raman and SERS recognition of β -carotene and haemoglobin fingerprints in human whole blood. *United States: Spectrochimica Acta Part A* 2011; 5: 915.
- [17] Uzunbajakava N, Lenfereink A, Kraan Y, Willekens B, Greve J, Otto C. Nonresonant Raman Imaging of Protein Distribution in Single Human Cells. *Luxembourg: Biopolymers* 2003; 72: 1.
- [18] Artemyev DN, Zakharov VP, Davydkin IL, Khristoforova JA, Lykina AA, Konyukhov VN, Kuzmina TP. Measurement of human serum albumin concentration using Raman spectroscopy setup. *Luxembourg: Opt. Quant Electron* 2016; 48: 337.