

The Effect of Noise in Raman Spectra on the Reconstruction of the Concentration of Amino Acids in the Mixture by Multivariate Curve Resolution (MCR) Analysis

Irina A. Matveeva*, Lyudmila A. Bratchenko, Oleg O. Myakinin, Elena N. Tupikova, and Valery P. Zakharov

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

* e-mail: m-irene-a@yandex.ru

Abstract. Changes in the concentration of free amino acids in biological tissues is a sign of impaired protein metabolism in patients with cancer. Recently, Raman spectroscopy has been used for early diagnostics of oncological diseases. The concentrations of individual components of biological tissue (for instance, the concentrations of amino acids) can be obtained by decomposing the tissue Raman spectrum. This study was designed to evaluate the effect of noise in the Raman spectra of individual amino acids on the result of the decomposition of the spectra of an amino acid mixture. As a decomposition method, we used Multivariate Curve Resolution-Alternating Least Squares (MCR-ALS) analysis and investigate experimental Raman spectra of amino acids and mathematically simulated Raman spectra of amino acid mixtures. Noise with different signal-to-noise ratios (SNR) was artificially added to both the experimental spectra of pure amino acids and the spectra of the mixtures. Concentration values for each amino acid obtained as a result of applying the MCR-ALS analysis have been compared with the corresponding true values and the correlation coefficients have been calculated. The results show a less pronounced negative effect of noise in the case when the spectra of pure amino acids (which were used as a basis for the MCR-ALS analysis) are noisy, and a more pronounced negative effect when the spectrum of the mixture is noisy. The accuracy of reconstruction of an amino acid is also negatively affected by strong background fluorescence in the amino acid spectrum. Moreover, the results indicate that using the basis spectra with a high SNR (SNR = 5) makes it possible to successfully estimate the amino acid concentrations in a mixture even when the Raman spectrum of the mixture is noisy and has a low SNR (SNR < 5). © 2021 Journal of Biomedical Photonics & Engineering.

Keywords: Signal-to-noise ratio; Raman spectrum; Raman scattering; multivariate curve resolution; free amino acids.

Paper #3424 received 23 May 2021; revised manuscript received 25 Jun 2021; accepted for publication 26 Jun 2021; published online 30 Jun 2021. [doi: 10.18287/JBPE21.07.020309](https://doi.org/10.18287/JBPE21.07.020309).

1 Introduction

Cancer is the second leading cause of death worldwide. The most common type of cancer was lung cancer. According to the World Health Organization, there was about 2.21 million new cases in 2020. Moreover, lung

cancer was the most common cause of cancer death in 2020 (1.80 million deaths). It is well known that lung cancer risk and mortality can be reduced by early detection of cases [1].

It is also known that during the malignant tumor development, changes in protein metabolism occur in the

patient's body. Many researchers have described changes in plasma free amino acid (PFAA) profiles in patients with cancer. Kubota, Meguid, and Hitch [2] analyzing PFAA in venous blood of patients with breast cancer, gastrointestinal tract cancer, and head and neck cancer suggested that PFAA profiles correlate diagnostically with the organ-site origin of three different kinds of malignant tumors. Miyagi et al. [3] determined the characteristics of the PFAA profiles in cancer patients with one of five types of cancer: lung, gastric, colorectal, breast, or prostate cancer. PFAA profiling for detecting lung cancer was also studied by Shingyoji et al. [4], Zhao et al. [5], and Proenza et al. [6]. These findings suggest that PFAA profiling has great potential for improving early detection of lung cancer.

It makes us look for new methods to analyze PFAA profiles. Recently, Raman spectroscopy has been used for early diagnostics of oncological diseases. Bratchenko et al. [7] have shown that this method can be used in the diagnosis of cancer, such as skin neoplasms. Moreover, Raman spectroscopy is an optical method that is relevant for the analysis of liquid media. So, we believe that Raman spectroscopy can be used for non-invasive analysis of blood plasma. The Raman spectra of PFAAs are specific and can be used to successfully evaluate PFAAs' concentration in a mixture of different substances by the Raman spectrum of this mixture.

In this study, we use Multivariate Curve Resolution–Alternating Least Squares (MCR–ALS) method to analyze the Raman spectra of amino acid mixtures. This method is widely used to reconstruct the concentration profiles of chemicals analysis [8, 9]. Recently, MCR–ALS method has found wide biological and medical applications [8, 10] and has been used for the analysis of spectral data when it is required to determine the concentrations of complex mixture components from spectra. For example, Xu and Rice [11] used a MCR spectral unmixing in fluorescence imaging. Chen et al. [12] used Raman spectroscopic detection of keratin with MCR analysis for automatic oral cancer diagnosis. Iwasaki et al. [13] investigated the possibilities of discrimination of breast cancer cells from normal mammary epithelial cells by Raman microspectroscopy and MCR analysis. It should be noted that the use of MCR–ALS analysis makes it possible not only to estimate the concentrations of components, but also to obtain their “pure” Raman spectra [14].

However, the MCR–ALS analysis can be sensitive to noise in the Raman spectra analyzed. In the practical application, the efficiency of evaluation of PFAAs' concentration in a mixture may decrease due to the fact that the Raman spectra contain a noise signal. We also suppose that, if MCR–ALS uses a known predetermined basis of amino acids, the result can be affected by noise in both the pure amino acid Raman spectra of the basis and Raman spectra of the mixtures. On the one hand, we can provide high-quality registration of basis Raman spectra (that is, the spectra of pure amino acids that are used in the MCR–ALS analysis) with high signal-to-noise ratio (SNR) using a spectroscopic setup with high

spectral resolution and increasing integration time. In addition, pure amino acids may be available for registration (or Raman microscopy of the samples may be used). And finally, it is enough to record the spectra (which are used as the basis) only once – then one can use them when analyzing other samples of mixtures. On the other hand, in a clinical setting, high-quality recording of Raman spectra can be difficult due to the large patient flow and limited time to examine a patient. Therefore, it can be assumed that the Raman spectra, which are subject to analysis by the MCR–ALS method, will have a lower SNR than the previously obtained Raman spectra of amino acids used as a basis for the MCR–ALS analysis.

This study was designed to evaluate an effect of the noise in Raman basis spectra of amino acids and Raman spectra of a mixture on their reconstruction from the mixture.

2 Materials and Methods

The experimental Raman spectra used in our study are recorded using a portable spectroscopic setup which includes a thermally stabilized LML-785.0RB-04 laser diode module as an excitation source (785 ± 0.1 nm central wavelength, 200 mW laser power) and a QE 65 Pro spectrometer (OceanOptics, Inc., USA) with CCD detector operating at -15 °C [15].

We used 20 standard proteinogenic amino acids (see Table 1 and Fig. 1). Amino acids are presented in crystalline powder form. The amino acids in the crystal form were placed onto the metal-coated slide. All spectra were registered at the room temperature. The registration of spectra using this system was carried out in 800–1000 nm with 0.2 nm spectral resolution that corresponds to the 240 – 2236 cm^{-1} . The Raman signal of amino acids was acquired from 3 accumulations each of 5 sec integration time. The mean Raman spectrum for each amino acid was averaged over three registered spectra.

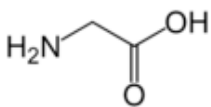
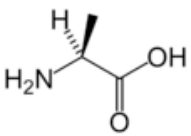
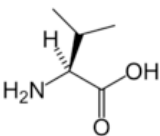
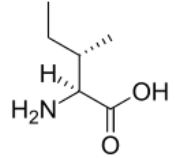
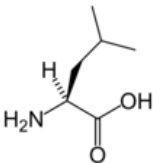
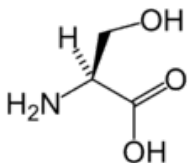
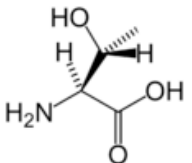
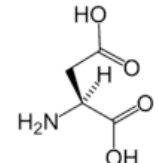
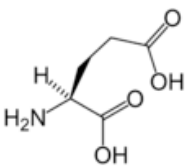
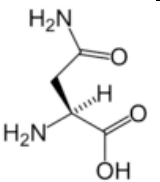
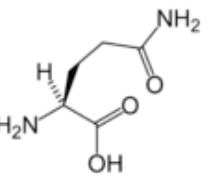
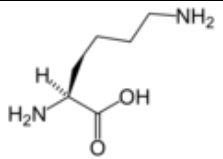
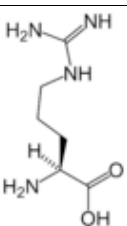
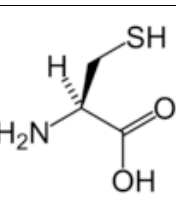
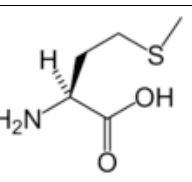
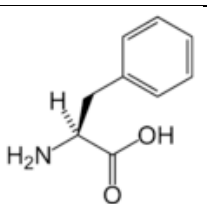
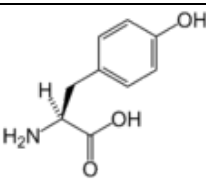
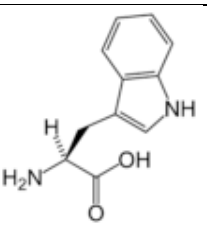
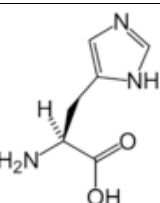
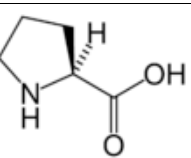
The preprocessing of the registered data includes only cosmic ray and dark noise removal that are automatically applied in the “Spectra Suite” software package (OceanOptics, Inc., USA) [16]. Examples of recorded and preprocessed spectra is shown in Fig. 1.

Then we have modeled amino acid mixtures. The “mixture” means the sum of 20 standard proteinogenic amino acids (see Table 1), taken in different quantities, that is, with different concentrations. The concentration of amino acids in the mixtures is chosen so that the mixtures correspond to real PFAA profiles of blood plasma samples studied by other researchers [3–6].

We define a Raman spectrum of an amino acid mixture as a mathematical sum of the products of the pure amino acid spectra of and its concentrations in the mixture:

$$S_M(\lambda) = c_{A1} \times s_{A1}(\lambda) + \dots + c_{AN} \times s_{AN}(\lambda), \quad (1)$$

Table 1 The standard proteinogenic amino acids.

Amino acid	Glycine	Alanine	Valine	Isoleucine
Letter code	Gly	Ala	Val	Ile
Formula	C ₂ H ₃ NO	C ₃ H ₅ NO	C ₅ H ₉ NO	C ₆ H ₁₁ NO
Skeleton formula				
Amino acid	Leucine	Serine	Threonine	Aspartic acid
Letter code	Leu	Ser	Thr	Asp
Formula	C ₆ H ₁₁ NO	C ₃ H ₅ NO ₂	C ₄ H ₇ NO ₂	C ₄ H ₅ NO ₃
Skeleton formula				
Amino acid	Glutamic acid	Asparagine	Glutamine	Lysine
Letter code	Glu	Asn	Gln	Lys
Formula	C ₅ H ₇ NO ₃	C ₄ H ₆ N ₂ O ₂	C ₅ H ₈ N ₂ O ₂	C ₆ H ₁₂ N ₂ O
Skeleton formula				
Amino acid	Arginine	Cysteine	Methionine	Phenylalanine
Letter code	Arg	Cys	Met	Phe
Formula	C ₆ H ₁₂ N ₄ O	C ₃ H ₅ NOS	C ₅ H ₉ NOS	C ₉ H ₉ NO
Skeleton formula				
Amino acid	Tyrosine	Tryptophan	Histidine	Proline
Letter code	Tyr	Trp	His	Pro
Formula	C ₉ H ₉ NO ₂	C ₁₁ H ₁₀ N ₂ O	C ₆ H ₇ N ₃ O	C ₅ H ₇ NO
Skeleton formula				

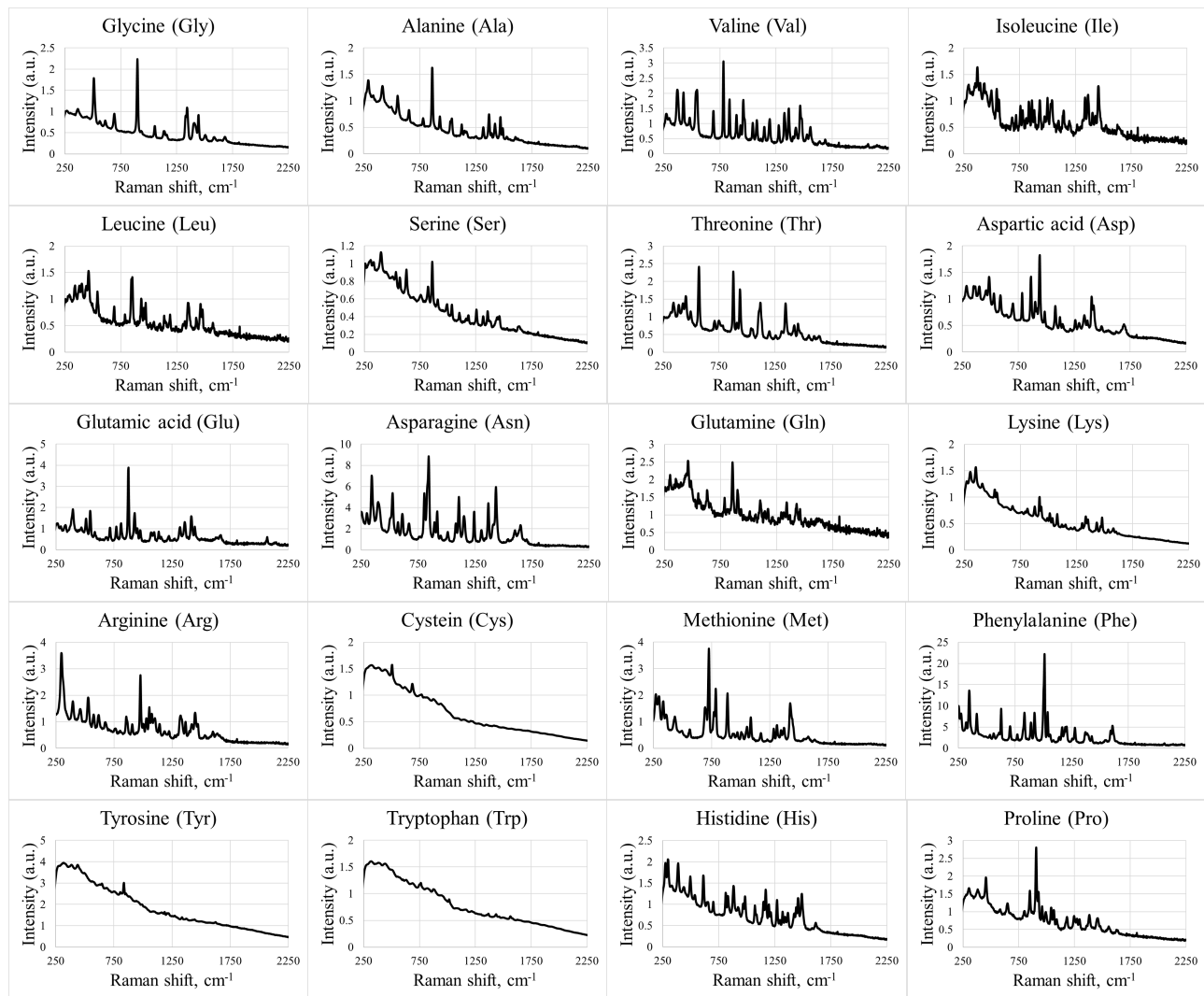


Fig. 1 Recorded and preprocessed Raman spectra of the amino acids.

where c_{A1}, \dots, c_{AN} – are the concentrations of pure amino acids in the mixture, $s_{A1}(\lambda), \dots, s_{AN}(\lambda)$ – are the pure amino acid spectra.

Using information on PFAA profiles in lung cancer patients [3–6], we have artificially modelled 10 Raman spectra: 5 spectra of lung cancer patient PFAA profiles and 5 spectra of control group PFAA profiles. Concentrations of amino acids in the mixtures are presented in Table 2.

The next step in our study was the simulation of a noise in the Raman spectra. It should be noted that the spectra recorded have a noise signal due to the spectroscopic setup. However, in this study, we investigate the effect of additively added noise. Therefore, the noise contained in the Raman spectra initially is not taken into account.

We added the noise as the random value process with hypothetically normal distribution, zero mean value and various standard deviations. In this case, the formula for a Raman spectrum looks like this:

$$S'(\lambda) = S(\lambda) + s_E(\lambda), \quad (2)$$

where $S(\lambda)$ is an original Raman spectrum, $s_E(\lambda)$ is the a noise spectrum.

To simulate different noise levels, we evaluated an additive noise level from a signal-to-noise ratio (SNR) metric proposed in [15]:

$$SNR = \frac{S}{6 \times \sigma_{s_E}}, \quad (3)$$

where S is the Raman signal level (intensity of Raman peak in 980–1025 cm^{-1} band); σ_{s_E} is the noise standard deviation.

To investigate the effect of the noise, it is necessary to compare different combinations of a noisy mixture and noisy basis spectra. For this purpose, we simulated Raman spectra of pure amino acids with SNRs equal to 1, 5, and 10 and Raman spectra of mixtures with SNRs equal to 2, 3, 4, and 5. Examples of the noisy spectra are shown in Fig. 2.

Table 2 Concentrations of amino acids in the mixtures.

Amino acids	Concentrations of amino acids, μM									
	Lung cancer patients					Control patients				
	№1 [4]	№2 [4]	№3 [5]	№4 [3]	№5 [6]	№1 [4]	№2 [4]	№3 [5]	№4 [3]	№5 [6]
Gly	214.3	208	298	220.7	133	209.7	203.5	454	207.2	120
Ala	324.2	366.5	643	356.1	213	343.1	353.5	611	348.1	231
Val	215.1	242.1	241	225.8	140	220.5	230	281	222.4	152
Ile	64.7	73.9	158	66.6	35.3	60.8	64.4	190	61.6	36.1
Leu	117.6	135.2	188	123.8	83.5	118.8	125.2	166	120.1	87.6
Ser	107.8	110.7	253	113.6	68.1	108.8	106.7	258	108.5	60.8
Thr	115.8	122.6	270	122.8	68.8	118.5	121.8	214	119.8	74.7
Asp	0	0	105	0	8.6	0	0	54	0	5.8
Glu	0	0	211	0	39.7	0	0	210	0	37.1
Asn	42.6	47.9	0	44.7	31.6	45.2	46.1	0	45.7	24
Gln	547.4	577.5	0	555.3	355	586.9	587.1	0	585.3	360
Lys	183.9	195.7	325	192.9	78.2	189.1	191.6	279	189.8	83.1
Arg	93.1	100.4	101	96.4	0	95.1	96.4	96	95.4	0
Cys	0	0	84	0	0	0	0	68	0	0
Met	24.2	26.9	57	25.6	16.1	25.8	26.8	61	26.1	14.8
Phe	59.5	66.5	139	61.3	40.3	59.6	61.3	109	59.5	35.4
Tyr	65.9	71.9	74	67.4	45.3	65.2	67.3	71	65.5	43.2
Trp	51.3	56.4	53	54.2	0	57	59.6	298	57.3	0
His	69.7	74.1	238	74.1	34.7	69.7	81.2	184	80.3	38.4
Pro	141.6	157.3	129	151.1	99.3	132.3	138.6	339	135.1	121

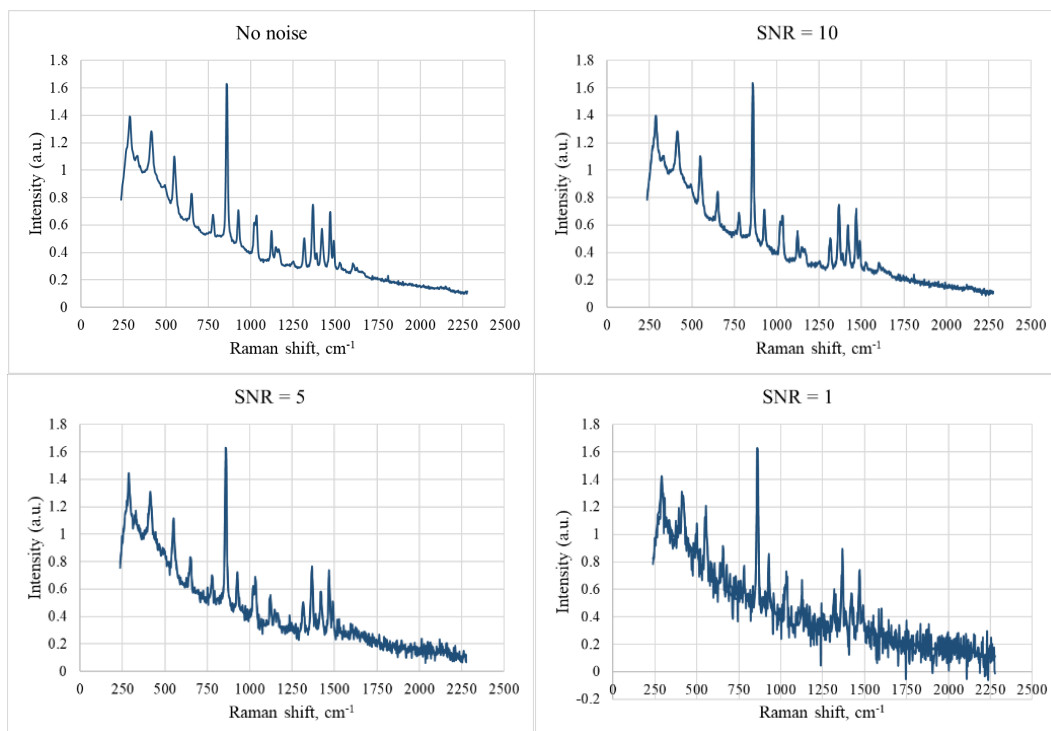


Fig. 2 Raman spectrum of alanine with different SNRs.

For unmixing spectra by MCR–ALS analysis we used a protocol by Felten et al. [14]. The main idea of MCR–ALS is to decompose the Raman spectra matrix D into smaller matrices C and S^T :

$$D = C \times S^T + E, \quad (4)$$

where C represents the concentration profiles for each of the amino acids, S^T is the pure amino acid spectra matrix, and E is the error matrix.

As a basis (S^T), we used the noisy amino acid spectra of 20 standard proteinogenic amino acids (see Table 1) to which we added noise in the previous step. After initial estimation is given for C , it is optimized iteratively using an alternative least squares algorithm (ALS) until convergence is reached [14].

3 Results and Discussion

In this study, we investigated the following combinations of a mixture and basis spectra:

- no noise in mixture spectra, no noise in basis spectra;
- no noise in mixture spectra, SNR = 10 for basis spectra;
- no noise in mixture spectra, SNR = 5 for basis spectra;
- no noise in mixture spectra, SNR = 1 for basis spectra;
- SNR = 10 for mixture spectra, no noise in basis spectra [17];

- SNR = 5 for mixture spectra, no noise in basis spectra [17];
- SNR = 1 for mixture spectra, no noise in basis spectra [17];
- SNR = 5 for mixture spectra, SNR = 5 for basis spectra;
- SNR = 4 for mixture spectra, SNR = 5 for basis spectra;
- SNR = 3 for mixture spectra, SNR = 5 for basis spectra;
- SNR = 2 for mixture spectra, SNR = 5 for basis spectra.

As a result of the MCR–ALS analysis, we have obtained a matrix of amino acid concentrations in the mixture spectra. Concentrations for each amino acid have been compared with the corresponding true values and the correlation coefficients have been calculated between the true concentration array and the obtained concentration array (see Tables 3, 4, 5).

Each concentration array corresponds to one of the amino acids and different mixtures (all the mixtures which we investigated). That is, each element of the array is a concentration of an amino acid in a mixture spectrum (a spectrum of one of the mixtures we used). The correlation coefficient indicates the degree of linear relation of the arrays and varies in the range between -1 and $+1$, where zero value corresponds to completely uncorrelated arrays.

Table 3 Correlation coefficients between obtained and true amino acid concentration values for the case of the noise added to basis spectra only.

Amino acids	SNR			
	∞ (no noise)	10	5	1
Gly	1.000	1.000	1.000	0.996
Ala	1.000	1.000	1.000	0.999
Val	1.000	1.000	1.000	0.998
Ile	1.000	0.999	1.000	0.988
Leu	1.000	0.997	0.999	0.908
Ser	1.000	0.999	1.000	0.994
Thr	1.000	1.000	1.000	0.999
Asp	1.000	0.997	1.000	0.996
Glu	1.000	1.000	1.000	1.000
Asn	1.000	1.000	1.000	0.999
Gln	1.000	1.000	1.000	1.000
Lys	1.000	1.000	0.999	0.982
Arg	1.000	1.000	1.000	0.994
Cys	1.000	0.987	0.999	0.998
Met	1.000	1.000	1.000	0.998
Phe	1.000	1.000	1.000	1.000
Tyr	1.000	0.988	0.979	0.989
Trp	1.000	0.999	0.999	0.996
His	1.000	1.000	1.000	0.998
Pro	1.000	0.999	1.000	1.000

Table 4 Correlation coefficients between obtained and true amino acid concentration values for the case of the noise added to mixture spectra only [17].

Amino acids	SNR			
	∞ (no noise)	10	5	1
Gly	1.000	0.999	0.991	0.966
Ala	1.000	0.999	0.995	0.959
Val	1.000	0.997	0.989	0.934
Ile	1.000	0.991	0.977	0.760
Leu	1.000	0.983	0.883	0.139
Ser	1.000	0.987	0.773	0.547
Thr	1.000	0.998	0.997	0.988
Asp	1.000	0.992	0.988	0.661
Glu	1.000	1.000	0.999	0.943
Asn	1.000	0.999	0.998	0.984
Gln	1.000	1.000	0.999	0.979
Lys	1.000	0.998	0.964	0.703
Arg	1.000	0.998	0.982	0.846
Cys	1.000	0.968	0.948	0.599
Met	1.000	0.997	0.953	-0.005
Phe	1.000	1.000	1.000	0.994
Tyr	1.000	0.923	0.448	0.344
Trp	1.000	0.991	0.966	0.641
His	1.000	0.999	0.990	0.901
Pro	1.000	0.998	0.992	0.899

Table 3 shows the correlation coefficients for the case when the basis spectra are noisy with a different SNR, and the mixtures have no noise. For comparison, Table 4 shows the results that we obtained in the previous study, when we investigated the effect of noise in the mixture Raman spectra on the quality of unmixing spectra [17].

As one can see from Table 3, the correlation coefficient between true and reconstructed concentrations of amino acids equals 1 for the case without noise. In the case of noisy Raman spectra, the quality of reconstruction of amino acid concentrations is expected to decrease. Nevertheless, in the cases of SNR from 10 to 1, the correlation coefficients are high for all amino acids and ranges from 0.90 to 1.

Paying attention to the comparison of these results with those obtained in the previous study (see Table 4), one can see the following. While in the case of a noise in the basis spectra, amino acids were reconstructed successfully with any of the considered SNRs, in the case of a noise in the mixture spectrum we failed to reconstruct all amino acids when SNR is less than 10. So, when SNR = 5, three out of 20 amino acids are reconstructed with an accuracy of less than 90%, and in the case of SNR = 1 there are already 11 such amino acids.

The results of our previous study [17] show that the concentrations of amino acids are restored with lower correlation coefficients if their Raman spectra have no intense peaks. That is, the ratio between Raman peaks and background, apparently caused by fluorescence, is

not high (compare the Raman spectra of cysteine, tyrosine, tryptophan with weak background fluorescence and the Raman spectra of glutamic acid, methionine, phenylalanine with strong background fluorescence in Fig. 1).

It can be concluded that a noise in the Raman spectrum of an amino acid mixture decreases the quality of reconstruction more than a noise in the basis, which is used in MCR-ALS. This may be due to the random nature of the noise we added to the spectra. It has zero mean value; therefore, during MCR-ALS analysis (see Eq. 4), the noise components compensate each other.

Table 5 demonstrates the correlation coefficients of reconstruction of mixture Raman spectra with different SNR using noisy basis Raman spectra of amino acids (SNR = 5). As expected, at higher mixture SNR (SNR > 4) the correlation coefficients are in a range from 0.9 to 1 for almost all amino acids except cysteine and tyrosine, which are characterized by low Raman peaks in all spectral range (see Fig. 1). For noisier mixture Raman spectra with SNR = 3 and SNR = 2 MCR-ALS method failed to reconstruct 5 and 6 amino acids with high quality, respectively. It should be noted that one of the failed components is cysteine, which either is absent in mixtures analyzed or its concentration is low.

As in our previous study [17], the concentrations of leucine, serine, cysteine, tyrosine, tryptophan are restored with lower correlation coefficients because random noise overlaps their spectra (which have a low ratio between Raman peaks and background fluorescence).

Table 5 Correlation coefficients between obtained and true amino acid concentration values for the case of the noise added to both basis spectra (SNR = 5 for all the cases) and mixture spectra.

Amino acids	SNR of mixture Raman spectra			
	5	4	3	2
Gly	0.996	0.995	0.988	0.991
Ala	0.994	0.990	0.988	0.956
Val	0.986	0.991	0.985	0.968
Ile	0.996	0.991	0.961	0.906
Leu	0.863	0.821	0.798	-0.085
Ser	0.876	0.869	0.628	0.571
Thr	0.992	0.986	0.985	0.949
Asp	0.983	0.947	0.941	0.964
Glu	0.998	0.999	0.997	0.996
Asn	0.997	0.983	0.991	0.971
Gln	0.999	1.000	0.997	0.995
Lys	0.969	0.939	0.928	0.776
Arg	0.992	0.993	0.987	0.941
Cys	0.555	0.650	0.727	0.403
Met	0.982	0.984	0.964	0.929
Phe	1.000	0.999	0.999	0.999
Tyr	0.451	0.666	0.214	0.589
Trp	0.907	0.903	0.876	0.829
His	0.991	0.992	0.966	0.978
Pro	0.991	0.992	0.984	0.947

It should be noted that the correlation coefficients for the cases of the basis spectra with noise and without noise are comparable. Thus, the approach proposed makes it possible to estimate amino acid concentrations from noisy Raman spectra of the mixture using basis spectra with noise, and this can be done almost as efficiently as if using pure basis spectra without noise.

Nevertheless, the presence of fluorescence in the spectra is not the only problem. The form of the spectrum, that is, the presence of Raman peaks in certain wavelength ranges, has a strong influence on the result of the decomposition of the mixture spectrum. For example, in the cases of SNR = 1 and SNR = 2, one can see extremely low negative correlation values for leucine (Leu) and methionine (Met) (see Tables 4 and 5). It can be explained by the fact that the Raman spectra of these amino acids do not contain peaks that are unique to only these amino acids; moreover, the spectra are noisy. Thus, the MCR method is not efficient enough in this case.

Having studied the effect of noise in Raman spectra on the reconstruction of the concentration of amino acids in the artificially modelled mixtures, we should check the obtained relations for real mixtures, when the noise is of different nature. It is very clear that the real mixture samples (i.e., biological tissue samples) may contain

other components such as lipids and proteins. In this case, the resolved spectra of the components may have some background contribution incorporated, a final estimation of component concentrations should be performed by fixing the pure spectra of the amino acids and leaving free the background contribution, which requires the correction of the protocol settings. In addition, the efficiency of reconstruction may be affected by removing background fluorescence and normalization of the basis spectra in order to equalize their intensities. Therefore, our future research will be devoted to experiments on real mixtures of amino acids and study of the influence of basis spectra preprocessing.

Disclosures

All authors declare that there is no conflict of interests in this paper.

Acknowledgments

This work was supported by the RFBR Grant № 19-52-18001 Bolg_a.

References

1. Cancer, World Health Organisation, 3 March 2021 (accessed 5 April 2021). [<https://www.who.int/news-room/fact-sheets/detail/cancer>].
2. A. Kubota, M. M. Meguid, and D. C. Hitch, “Amino acid profiles correlate diagnostically with organ site in three kinds of malignant tumors,” *Cancer* 69(9), 2343–2348 (1992).
3. Y. Miyagi, M. Higashiyama, A. Gochi, M. Akaike, T. Ishikawa, T. Miura, and M. Moriyama, “Plasma free amino acid profiling of five types of cancer patients and its application for early detection,” *PloS One* 6(9), e24143 (2011).
4. M. Shingyoji, T. Iizasa, M. Higashiyama, F. Imamura, N. Saruki, A. Imaizumi, and M. Yamakado, “The significance and robustness of a plasma free amino acid (PFAA) profile-based multiplex function for detecting lung cancer,” *BMC Cancer* 13(1), 77 (2013).
5. Q. H. Zhao, Y. Cao, Y. Wang, C. L. Hu, A. L. Hu, L. Ruan, and M. Ren, “Plasma and tissue free amino acid profiles and their concentration correlation in patients with lung cancer,” *Asia Pacific journal of clinical nutrition* 23(3), 429–436 (2014).
6. A. M. Proenza, J. Oliver, A. Palou, and P. Roca, “Breast and lung cancer are associated with a decrease in blood cell amino acid content,” *The Journal of Nutritional Biochemistry* 14(3), 133–138 (2003).
7. I. A. Bratchenko, L. A. Bratchenko, A. A., Moryatov, Y. A., Khristoforova, D. N. Artemyev, O. O. Myakinin, A. E. Orlov, S. V. Kozlov, and V. P. Zakharov, “In vivo diagnosis of skin cancer with a portable Raman spectroscopic device,” *Experimental Dermatology* 30(5), 652–663 (2021).
8. A. de Juan, R. Tauler, “Multivariate Curve Resolution: 50 years addressing the mixture analysis problem – A review,” *Analytica Chimica Acta* 1145(8), 59–78 (2020).
9. M. Garrido, F. X. Rius, and M. S. Larrechi, “Multivariate curve resolution–alternating least squares (MCR-ALS) applied to spectroscopic data from monitoring chemical reactions processes,” *Analytical and Bioanalytical Chemistry* 390(8), 2059–2066 (2008).
10. H. Noothalapati, K. Iwasaki, and T. Yamamoto, “Biological and medical applications of multivariate curve resolution assisted Raman spectroscopy,” *Analytical Sciences* 33(1), 15–22 (2017).
11. H. Xu, B. W. Rice, “In-vivo fluorescence imaging with a multivariate curve resolution spectral unmixing technique,” *Journal of Biomedical Optics* 14(6), 064011 (2009).
12. P. H. Chen, R. Shimada, S. Yabumoto, H. Okajima, M. Ando, C. T. Chang, L. T. Lee, Y. K. Wong, A. Chiou, and H. O. Hamaguchi, “Automatic and objective oral cancer diagnosis by Raman spectroscopic detection of keratin with multivariate curve resolution analysis,” *Scientific Reports* 6(1), 1–9 (2016).
13. K. Iwasaki, A. Araki, C. M. Krishna, R. Maruyama, T. Yamamoto, and H. Noothalapati, “Identification of Molecular Basis for Objective Discrimination of Breast Cancer Cells (MCF-7) from Normal Human Mammary Epithelial Cells by Raman Microspectroscopy and Multivariate Curve Resolution Analysis,” *International Journal of Molecular Sciences* 22(2), 800 (2021).
14. J. Felten, H. Hall, J. Jaumot, R. Tauler, A. De Juan, and A. Gorzsás, “Vibrational spectroscopic image analysis of biological material using multivariate curve resolution–alternating least squares (MCR-ALS),” *Nature protocols* 10(2), 217–240 (2015).
15. L. A. Bratchenko, I. A. Bratchenko, A. A. Lykina, M. V. Komarova, D. N. Artemyev, O. O. Myakinin, A. A. Moryatov, I. L. Davydkin, S. V. Kozlov, and V. P. Zakharov, “Comparative study of multivariate analysis methods of blood Raman spectra classification,” *Journal of Raman Spectroscopy* 51(2), 279–292 (2020).
16. SpectraSuite, Ocean Insight (accessed 29 July 2020). [<https://oceanoptics.ru/spectrometers/59-spectrasuite.html>].
17. I. A. Matveeva, O. O. Myakinin, Y. A. Khristoforova, I. A. Bratchenko, E. N. Tupikova, and V. P. Zakharov, “Possibilities for decomposing Raman spectra of amino acids mixture by Multivariate Curve Resolution (MCR) analysis,” *SPIE Proceedings* 11582, 115821G (2020).