

Analytical Methods

Accepted Manuscript

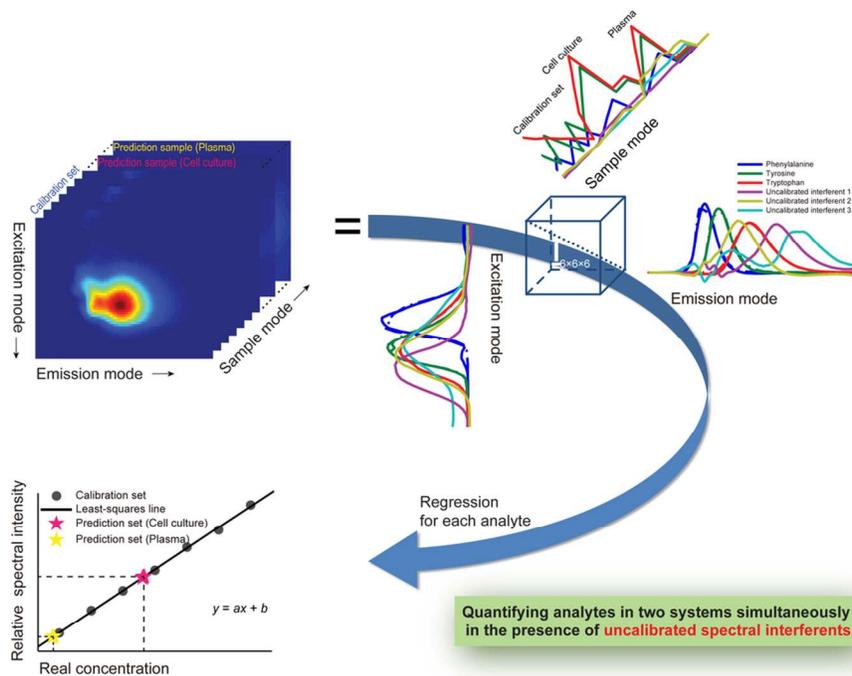


This is an *Accepted Manuscript*, which has been through the Royal Society of Chemistry peer review process and has been accepted for publication.

Accepted Manuscripts are published online shortly after acceptance, before technical editing, formatting and proof reading. Using this free service, authors can make their results available to the community, in citable form, before we publish the edited article. We will replace this *Accepted Manuscript* with the edited and formatted *Advance Article* as soon as it is available.

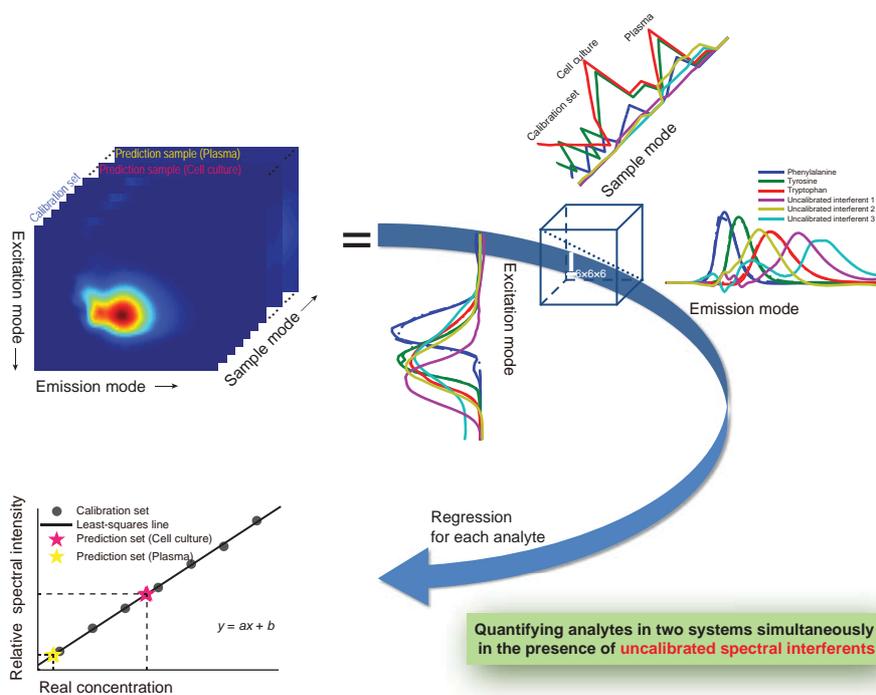
You can find more information about *Accepted Manuscripts* in the [Information for Authors](#).

Please note that technical editing may introduce minor changes to the text and/or graphics, which may alter content. The journal's standard [Terms & Conditions](#) and the [Ethical guidelines](#) still apply. In no event shall the Royal Society of Chemistry be held responsible for any errors or omissions in this *Accepted Manuscript* or any consequences arising from the use of any information it contains.



An efficient method, based on three-way calibration coupled with EEMs, is proposed for determining aromatic amino acids in different systems simultaneously, even in the presence of three uncalibrated interferences.

103x78mm (300 x 300 DPI)



Graphical abstract:

An efficient method, based on three-way calibration coupled with EEMs, is proposed for determining aromatic amino acids in different systems simultaneously, even in the presence of three uncalibrated interferences.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Simultaneous determination of aromatic amino acids in different systems using three-way calibration based on the PARAFAC-ALS algorithm coupled with EEM fluorescence: Exploration of second-order advantages

Chao Kang, Hai-Long Wu^{*a}, Shou-Xia Xiang, Li-Xia Xie, Ya-Juan Liu, Yong-Jie Yu, Jing-Jing Sun, Ru-Qin Yu

Received (in XXX, XXX) Xth XXXXXXXXXX 20XX, Accepted Xth XXXXXXXXXX 20XX

DOI: 10.1039/b000000x

Abstract: A practical analytical method based on intrinsic fluorescence is proposed for simultaneous determination of L-phenylalanine, L-tyrosine, and L-tryptophan in cell culture and human plasma. By using three-way calibration method coupled with excitation-emission matrix fluorescence, the proposed method successfully achieved quantitative analysis of the three aromatic amino acids in the two different complex systems simultaneously, even in the presence of three unknown, uncalibrated serious interferents. The method needs little preparation by using “mathematical separation” instead of chemical or physical separation, which makes it efficient and environmentally friendly. Satisfactory results have been achieved for calibration, validation, and prediction sets. For phenylalanine, tyrosine, and tryptophan, the calibration ranges are 6.00 to 60.00, 0.40 to 4.00, and 0.10 to 1.00 $\mu\text{g mL}^{-1}$ respectively. The average spike recoveries (mean \pm standard deviation) are $98.5 \pm 7.8\%$, $103.7 \pm 6.9\%$, and $102.3 \pm 7.9\%$ respectively. The relative errors are -4.2% , 6.3% , and -0.8% for real contents of phenylalanine, tyrosine, and tryptophan in cell culture respectively. Additionally, we discussed the potentiality of three-way calibration method for determining analytes of interest in different systems simultaneously, to further explore the second-order advantages. The paired *t*-test results indicate that the predicted results between predicting in two systems simultaneously and predicting in single system individually have no significant difference. The satisfactory results obtained in this work indicate that the use of three-way calibration method coupled with EEMs array is a promising tool for multi-component simultaneous determination in multiple complex systems containing uncalibrated spectral interferents.

1 Introduction

L-phenylalanine, L-tyrosine, and L-tryptophan are three of the building blocks of polypeptides and proteins. They participate in many functions of the living cell, such as signal transduction¹. As important small molecules, they also participate in metabolism, such as citric acid cycle. Both high and low levels of phenylalanine, tyrosine, and tryptophan can influence the normal functions and metabolism of the body. Therefore, quantitative analysis of phenylalanine, tyrosine, and tryptophan in complex biological systems is of great interest. Many analytical methods have been reported²⁻⁵. The most frequently used method is (liquid or gas) chromatography combined with mass spectroscopy. However, these analytical separations usually involve careful sample pretreatment or separation steps for the determination in complex biological systems. Additionally, the step of sample pretreatment always induces bias to the predicted concentration level of analyte.

One alternative to chromatographic methods is fluorescence spectroscopy. Fluorescence measurement can be applied to

dissolve a wide range of problems in the chemical and biological sciences. Considering that phenylalanine, tyrosine, and tryptophan are natural fluorophores, fluorescence spectroscopy is potential for quantitative analysis of them. Fluorescence measurement is highly sensitive⁶⁻¹⁰, but fluorescence detection cannot provide high selectivity¹⁰. For quantitative analysis of phenylalanine, tyrosine, and tryptophan in complex system, peculiar situations exist: using fluorescence signal at the maximum emission wavelength, one can only determine one analyte at a time with purifying it from the other two analytes and plasma background, since the classical one-way (zero-order) calibration method^{11,12} requires signal must be fully selective for the analyte of interest. With two-way (first-order) multivariate calibration method¹¹⁻¹⁷, for example, based on fluorescence emission spectra, the three analytes can be determined simultaneously. It makes sense to treat multivariate measurements simultaneously in data analysis, by which we could obtain more information in the form of correlation^{18,19}. However, the spectrum for the analyte of interest must be partially different from the spectra of all other responding species. Additionally, calibration standards must be representative of the

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

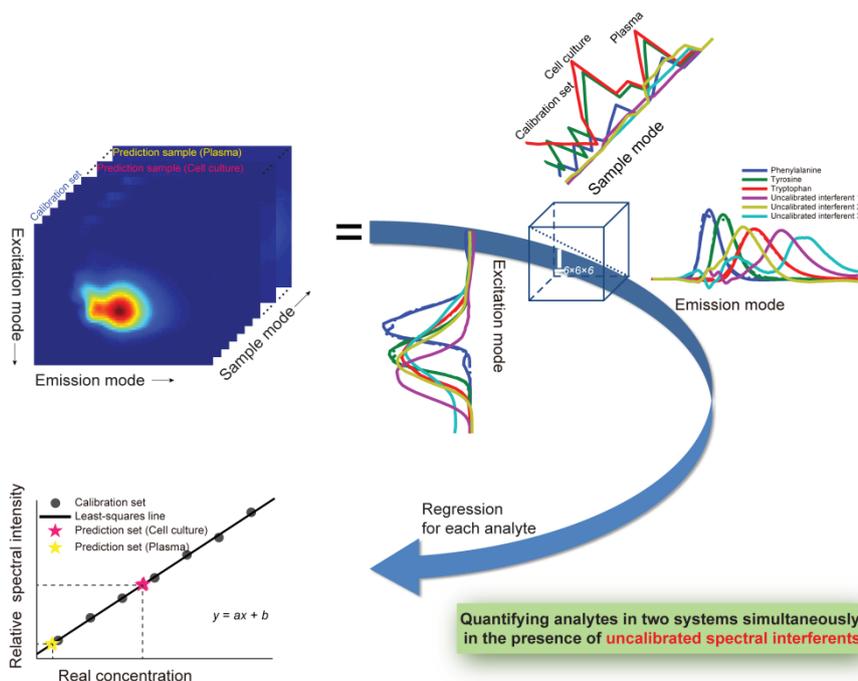


Fig. 1 Illustration of the assay.

samples containing any interfering species in the system studied. As the rapid development of fluorescence instrument such as excitation-emission matrix (EEM) fluorescence spectroscopy^{20, 21}, this prerequisite can be overcome through three-way (second-order) calibration method²¹⁻³². This is due to the second-order advantage^{11, 33-35}, which means analytes can be quantitatively analyzed even in the presence of uncalibrated interferent. One of

phenylalanine, tyrosine, and tryptophan respectively, which give accurate prediction for calibration set, validation set and real contents of analytes in different biofluids (see Figure 1). In addition, we discuss the ability of three-way calibration to predict analytes of interest in different complex systems simultaneously, to further explore the second-order advantages.

2 Theory

2.1 Trilinear component model

In three-way (second-order) calibration, one of the most commonly used models is the trilinear component model (also known as PARAFAC/CANDECOMP model)³⁸⁻⁴³, which is proposed by Harshman³⁶ and Carroll and Chang³⁷ independently. When an EEM of a sample is measured at I excitation wavelengths and J emission wavelengths, the collection of EEMs of calibration samples and prediction samples forms a three-way data array \mathbf{X} dimensioned I excitation wavelengths by J emission wavelengths by K samples, in which each element x_{ijk} can be expressed as follows:

$$x_{ijk} = \sum_{n=1}^N a_{in} b_{jn} c_{kn} + e_{ijk}$$

$$\text{for } i = 1, 2, \dots, I; j = 1, 2, \dots, J; k = 1, 2, \dots, K. \quad (1)$$

where a_{in} , b_{jn} , and c_{kn} correspond to the underlying excitation spectra matrix $\mathbf{A}_{I \times N}$, emission spectra matrix $\mathbf{B}_{J \times N}$, and relative

spectral intensity matrix $\mathbf{C}_{K \times N}$, respectively, where N represents the detectable species in the system. The term e_{ijk} is the element of the three-way residual data array $\mathbf{E}_{I \times J \times K}$. Regardless of scaling and permutation, the decomposition of the trilinear component model will be a unique one given that $k_A + k_B + k_C \geq 2F + 2$, where k_A , k_B , and k_C are the k -ranks of the profile matrices \mathbf{A} , \mathbf{B} , and \mathbf{C} respectively⁴⁴.

In addition, the trilinear component model can be expressed as the following stretched matrix forms:

$$\mathbf{X}_{I \times JK} = \mathbf{A}(\mathbf{C} \odot \mathbf{B})^T + \mathbf{E}_{I \times JK} \quad (2)$$

$$\mathbf{X}_{J \times KI} = \mathbf{B}(\mathbf{A} \odot \mathbf{C})^T + \mathbf{E}_{J \times KI} \quad (3)$$

$$\mathbf{X}_{K \times IJ} = \mathbf{C}(\mathbf{B} \odot \mathbf{A})^T + \mathbf{E}_{K \times IJ} \quad (4)$$

where \odot indicates the Khatri-Rao product. Provided that matrices $\mathbf{A} \in \mathbb{R}^{I \times N}$ and $\mathbf{B} \in \mathbb{R}^{J \times N}$, their Khatri-Rao product is a matrix of size $(IJ) \times N$ and defined by

$$\mathbf{A} \odot \mathbf{B} = \begin{bmatrix} a_{11}\mathbf{b}_1 & a_{12}\mathbf{b}_2 & \dots & a_{1N}\mathbf{b}_N \\ a_{21}\mathbf{b}_1 & a_{22}\mathbf{b}_2 & \dots & a_{2N}\mathbf{b}_N \\ \vdots & \vdots & \ddots & \vdots \\ a_{I1}\mathbf{b}_1 & a_{I2}\mathbf{b}_2 & \dots & a_{IN}\mathbf{b}_N \end{bmatrix}$$

2.2 PARAFAC-ALS method

In general, the PARAFAC-ALS algorithm^{21, 26, 36} is carried out by using alternating least squares principle. According to equations (2), (3), and (4), the solutions in the three modes can be obtained as follows

$$\mathbf{A} = \mathbf{X}_{I \times JK}((\mathbf{C} \odot \mathbf{B})^T)^+ \quad (5)$$

$$\mathbf{B} = \mathbf{X}_{J \times KI}((\mathbf{A} \odot \mathbf{C})^T)^+ \quad (6)$$

$$\mathbf{C} = \mathbf{X}_{K \times IJ}((\mathbf{B} \odot \mathbf{A})^T)^+ \quad (7)$$

Through the decomposition of three-way array by the PARAFAC-ALS algorithm, the relative profiles of each mode can be obtained. The decomposition of the trilinear component model joins the calibration set together with prediction set, thus concentration information can be obtained in a separate univariate regression step. In this work, we take the real concentration as the independent variable. The univariate regression is expressed as

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_P \end{bmatrix} = \begin{bmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_P \end{bmatrix} \begin{bmatrix} b_0 \\ b_1 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_P \end{bmatrix} \quad \text{or } \mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e} \quad (8)$$

where P is the number of calibration samples. The model parameter is estimated by $(\bar{b}_0, \bar{b}_1)^T = \mathbf{X}^+ \mathbf{y}$. Then the analyte concentration is predicted by $\bar{x}_{\text{unk}} = (y_{\text{unk}} - \bar{b}_0) / \bar{b}_1$ for an unknown sample, where y_{unk} represents the decomposed relative spectral intensity of the analyte.

In this study, we use random initialization to start the iterative optimizing procedures of the PARAFAC-ALS algorithm. The optimizing procedures terminated when the following criterion is satisfied (in this study, we set the threshold $\varepsilon = 1 \times 10^{-9}$).

$$\left| \frac{\text{SSR}^{(m)} - \text{SSR}^{(m-1)}}{\text{SSR}^{(m-1)}} \right| < \varepsilon \quad (9)$$

where SSR is residual sum of squares, $\text{SSR} = \sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K e_{ijk}^2$, and m is the current number of iterations. A maximal iteration number (10000) is adopted to avoid possible excess slow convergence.

2.3 Figures of merit

Figures of merit are analytical parameters used for evaluating performance of calibration method. Different approaches have been discussed in the literature for computing figures of merit for higher-order methodologies^{45, 46}.

The root mean square error of prediction (RMSEP)^{12, 18} is computed by

$$\text{RMSEP} = \sqrt{\frac{1}{P} \sum_{p=1}^P (y_p - \bar{y}_p)^2} \quad (10)$$

where P is number of samples, y_p and \bar{y}_p are the actual and predicted concentration respectively.

As regards the sensitivity, the following equation⁴⁵ seems to apply to the present data:

$$\text{SEN}_n = k_n / \sqrt{\left(\left((\mathbf{A}_{\text{ex}}^T (\mathbf{I} - \mathbf{A}_{\text{un}} \mathbf{A}_{\text{un}}^+) \mathbf{A}_{\text{ex}}) * (\mathbf{B}_{\text{ex}}^T (\mathbf{I} - \mathbf{B}_{\text{un}} \mathbf{B}_{\text{un}}^+) \mathbf{B}_{\text{ex}}) \right)^{-1} \right)_{nn}} \quad (11)$$

where k_n is the pure analyte signal at unit concentration, \mathbf{A}_{ex} and \mathbf{B}_{ex} are the matrices containing the profiles for all expected components in each mode, \mathbf{A}_{un} and \mathbf{B}_{un} are the matrices containing the profiles for all unexpected components in each mode, "*" represents the Hadamard product, and (n, n) element corresponds to the n th analyte of interest.

The limit of detection (LOD) and limit of quantitation (LOQ) are estimated⁴⁶ as follows:

$$\text{LOD} = 3.3 \times s(0) \quad (12)$$

$$\text{LOQ} = 10 \times s(0) \quad (13)$$

where $s(0)$ is the standard error in the predicted concentration for the method blank samples, computed by

$$s(0) = \sqrt{h_0 s_c^2 + h_0 \frac{s_x^2}{\text{SEN}^2} + \frac{s_x^2}{\text{SEN}^2}} \quad (14)$$

where h_0 is the method blank sample leverage, s_c^2 is the variance in calibration concentrations, s_x^2 is the variance in the instrumental signal, and SEN is the analyte sensitivity.

The spike recovery⁴⁷ is computed as

$$\% \text{ spike recovery} = \frac{C_{\text{spiked sample}} - C_{\text{unspiked sample}}}{C_{\text{added}}} \times 100 \quad (15)$$

where C stands for concentration.

3 Experimental

3.1 Apparatus and software

Fluorescence spectral measurements were performed on an F-7000 fluorescence spectrophotometer (HITACHI, Tokyo, Japan) equipped with a Xenon lamp. All measurements were recorded in a 10 mm quartz cell at room temperature. The EEMs were recorded at emission wavelengths between 202 and 500 nm (every 2 nm), at different excitation wavelengths between 202 nm and 320 nm (every 2 nm). Excitation and emission slit-widths were both set to be 5 nm, scan speed was set at 12,000 nm min⁻¹, and detector voltage was 500 V.

The routines used were written in the MATLAB environment and run on a 3.07 GHz Intel (R) Core (TM) i7 CPU with 3 GB RAM under Windows 7 operating system.

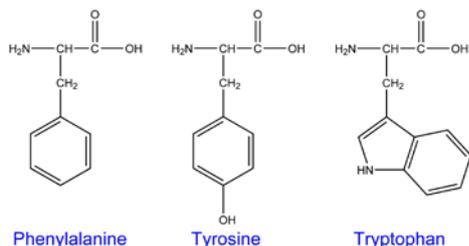


Fig. 2 Chemical structures of phenylalanine, tyrosine, and tryptophan.

3.2 Reagents and chemicals

L-phenylalanine (99%), L-tyrosine (99%), and L-tryptophan (99%) (the structures of the three amino acids are shown in [Figure 2](#)) were purchased from Aladdin (Aladdin, Shanghai, China). All other chemicals used were of analytical reagent grade. Ultrapure water was produced by the Milli-Q Gradient A10 system (Millipore, Billerica, USA). The cell culture RPMI-1640 was purchased from Thermo Scientific (Thermo Scientific, Beijing, China), the human plasma from YuanHengJinMa Biotechnology Development Co. (YuanHengJinMa, Beijing, China). The buffer is $\text{Na}_2\text{HPO}_4\text{-C}_6\text{H}_8\text{O}_7$ aqueous (pH 2.2). Individual stock solutions were prepared in 1.0 % formic acid aqueous (Phenylalanine, tyrosine, and tryptophan at concentrations of about 5000, 1000, 1000 $\mu\text{g mL}^{-1}$ respectively). All the solutions were stored in the refrigerator at 4°C.

3.3 Three-way data array

The calibration set consists of seven samples, in which the seven concentration levels of the three analytes of interest are designed by a $U_7(7^3)$ uniform design^{48, 49}, since the uniform design could provide an appropriate experimental design for this seven-level, three-factor concentration design with only seven experiments. For preparing a given calibration sample, the analytes were mixed in the volumetric flask, by taking appropriate volumes of phenylalanine, tyrosine, and tryptophan stock solutions and diluting to 10.00 mL with buffer. The validation set 1 consists of five spiked samples. Each of them was prepared as follows: volume of 1.00 mL cell culture was spiked with specific contents of all analytes (selected from their corresponding calibration ranges), and diluted to 10.00 mL with buffer. The prediction set 1 consists of three cell culture samples, which were used to determine the contents of phenylalanine, tyrosine, and tryptophan in cell culture RPMI-1640. The validation set 2 and prediction set 2 for human plasma were prepared similarly. Three method blank samples were prepared. In addition, three reference samples were prepared to provide the reference spectra of analytes of interest. The details on the concentration designs of these samples are given in [Table 1](#). Then, an excitation-emission-sample three-way data array (60×150×26) could be constructed.

Table 1 The concentration designs of phenylalanine, tyrosine, and tryptophan.

Sample	Concentrations ($\mu\text{g mL}^{-1}$)			Cell culture (mL/10 mL)	Plasma ($\mu\text{L}/10\text{ mL}$)
	Phenylalanine	Tyrosine	Tryptophan		
Calibration set					
cal01	24.00	2.80	1.00		
cal02	51.00	1.00	0.85		
cal03	6.00	4.00	0.70		
cal04	33.00	2.20	0.55		
cal05	60.00	0.40	0.40		
cal06	15.00	3.40	0.25		
cal07	42.00	1.60	0.10		
Validation set 1					
val01	6.00	0.40	0.10	1.00	
val02	19.50	1.30	0.32	1.00	
val03	33.00	2.20	0.55	1.00	
val04	46.50	3.10	0.77	1.00	
val05	60.00	4.00	1.00	1.00	
Prediction set 1					
pre01				1.00	
pre02				1.00	
pre03				1.00	
Validation set 2					
val06	6.00	0.40	0.10		2.50
val07	19.50	1.30	0.32		2.50
val08	33.00	2.20	0.55		2.50
val09	46.50	3.10	0.77		2.50
val10	60.00	4.00	1.00		2.50
Prediction set 2					
pre04					2.50
pre05					2.50
pre06					2.50

4. Results and Discussion

4.1 Spectral properties of the analytes and the biofluids

[Figure 3](#) shows the three-dimensional landscapes of EEM fluorescence for calibration sample cal04, prediction sample pre02 (Cell culture), pre05 (Human plasma), and method blank sample bla02. From the landscape of cal04, one can see that the spectra of phenylalanine, tyrosine, and tryptophan overlap seriously. For pre02, only one peak can be seen, although the cell culture is so complex that it may contains many fluorescence responsive constituents, there is no resolution among them. So does the human plasma. The spectra of analytes totally overlap with that of both biofluids.

In all samples severe Rayleigh and Raman scattering is present. These nonlinear factors can lead the EEMs array to deviate the trilinear component model, which is a prerequisite for the PARAFAC-ALS algorithm to decompose profile matrices correctly. We have handled scattering using interpolation⁵⁰ in the areas affected by first- and second-order Rayleigh and Raman scattering. The width is a two-element vector defining how many nanometres to the left and right of the scatter centre is removed, of which first-order Rayleigh width = [10, 10], Raman width = [8, 8], and second-order Rayleigh width = [15, 15].

To exploring the linear range of each analyte, a series of pure standards was prepared for each analyte individually. For analyte phenylalanine, tyrosine, and tryptophan, the linear ranges are 3.00 to 120.00 ($R^2 = 0.999$), 0.20 to 8.00 ($R^2 = 0.999$), and 0.05 to 2.00

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

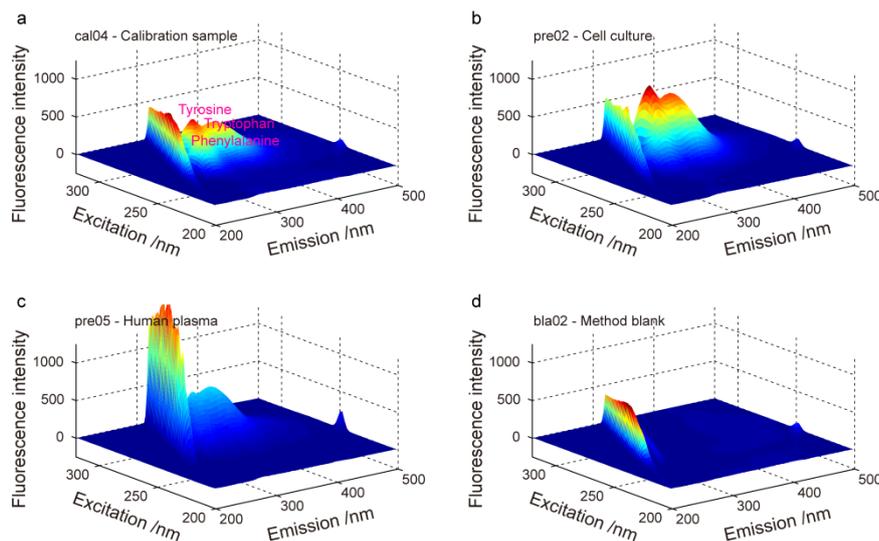


Fig. 3 Three-dimensional landscapes of EEMs of sample cal04, pre02 pre05 and bla02.

($R^2 = 0.999$) $\mu\text{g mL}^{-1}$ respectively. The selected calibration ranges of phenylalanine, tyrosine, and tryptophan are 6.00 to 60.00, 0.40 to 4.00, and 0.10 to 1.00 $\mu\text{g mL}^{-1}$ respectively.

4.2 Estimating the number of components

For quantitative analysis of the three aromatic amino acids based on EEM fluorescence, one needs to extract the pure profiles of each analyte in the seriously overlapped peak among analytes and interferents by chemical or physical separation or mathematical separation, then to predict the real concentrations of them. Three-way calibration method might be the right one for mathematical separation because of its “second-order advantage”. Herein we will quantitatively analyze aromatic amino acids in these two different biofluids simultaneously using three-way calibration coupled with intrinsic EEM fluorescence and further explore the second-order advantages.

The number of components for the trilinear component model should be estimated before the application of the PARAFAC-ALS algorithm^{36, 37, 51, 52}. We used the core consistency diagnostic method⁵² to select the number of components. It is given as the percentage of variation in a Tucker3 core array consistent with the theoretical superidentity array. When the core consistency drops from a high value to a low value, it indicates that an appropriate number of components have been attained.

By joining the EEMs of the seven calibration samples, five cell culture spiked validation samples, three cell culture prediction samples, five human plasma spiked validation samples, three human plasma prediction samples and three method blank samples together, a three-way data array of $60 \times 150 \times 26$ is

constructed. The background was subtracted by the method blank samples. For this EEMs three-way data array, the core consistency values are 100.0 %, 100.0 %, 99.2 %, 60.7 %, 43.1 %, 39.2 % and 1.0 % when $N = 1, 2, 3, 4, 5, 6$ and 7 respectively. Considering that rank determination is not always straightforward, especially in the presence of noise or in complex system¹⁸, also the chemical criteria (e.g. reference spectrum) and the root mean square errors of calibration (RMSECs) are used to evaluate the appropriate number of components. Finally, $N = 6$ is selected as the optimum number of components for the EEMs three-way data array.

4.3 Testing the calibration model internally

Prior to the following analysis, we perform the PARAFAC-ALS method ($N = 3$) on the calibration set to testing the calibration model internally. For analytes of interest phenylalanine, tyrosine, and tryptophan, RMSECs are 0.51, 0.03 and 0.01 $\mu\text{g mL}^{-1}$ respectively. The average recoveries (mean \pm standard deviation) of the calibration set are $99.0 \pm 4.7\%$, $99.5 \pm 2.4\%$ and $99.8 \pm 1.6\%$ respectively. The explained variance is 99.9%. All these results distinctly indicate that the trilinear component model constructing on EEM fluorescence data is quite probably a good one.

4.4 Prediction for validation sets and prediction sets

The validation sets are used to prove that the PARAFAC-ALS method can accurately predict the contents of phenylalanine, tyrosine, and tryptophan even in the presence of uncalibrated interferents in cell culture and human plasma simultaneously.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

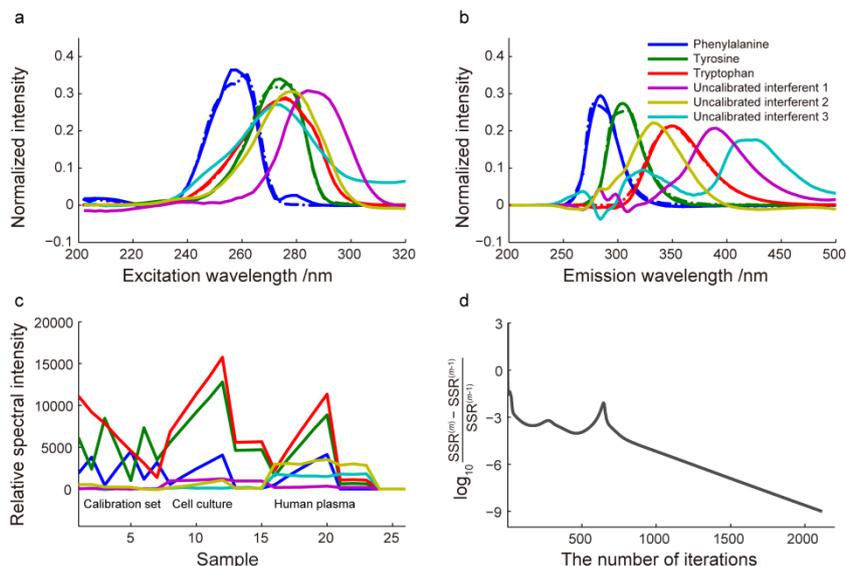


Fig. 4 Actual profiles (dash-dot lines) and decomposed profiles (solid lines) in each mode using the PARAFAC-ALS algorithm ($N = 6$).

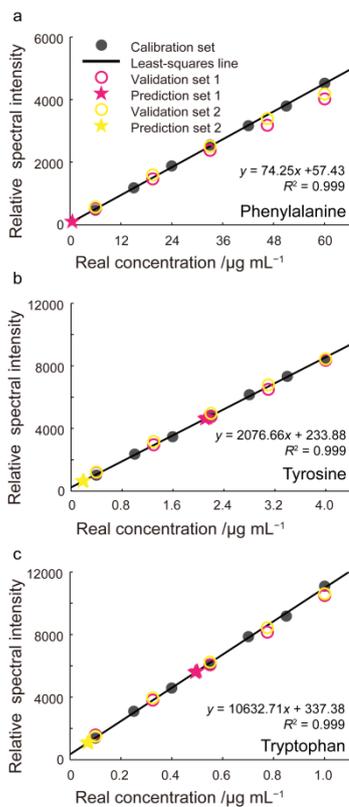


Fig. 5 Classical univariate regression of decomposed relative spectral intensity against real concentrations for phenylalanine, tyrosine, and tryptophan respectively.

Prediction sets are used to determine the real contents of phenylalanine, tyrosine, and tryptophan in cell culture and human plasma. We used the PARAFAC-ALS algorithm with $N = 6$ to decompose the trilinear component model, followed by linear classical univariate regression for each analyte. With $N = 6$ for the EEMs data array, the PARAFAC-ALS algorithm accurately decomposed the excitation spectra matrix $\mathbf{A}_{60 \times 6}$, emission spectra matrix $\mathbf{B}_{150 \times 6}$, and relative spectral intensity matrix $\mathbf{C}_{26 \times 6}$. The explained variance is 99.9% and the number of iterations is 2115. Figure 4 shows the actual profiles (dash-dot lines) and decomposed profiles (solid lines) in each mode using the PARAFAC-ALS algorithm ($N = 6$). From the decomposed profiles in each mode, one can see that this is indeed a high collinear and complex system. The decomposed excitation and emission profiles of each analyte are in good agreement with the actual profiles based on pure analyte standards. Three uncalibrated interferents were resolved. Uncalibrated interferent 1 comes from cell culture and, uncalibrated interferents 2 and 3 come from human plasma. The fluorescence excitation and emission spectra of all interferents overlap with that of analytes.

Figure 5 shows the linear classical univariate regression of decomposed relative spectral intensity against real concentration for phenylalanine, tyrosine, and tryptophan respectively. The spiked validation samples confirmed that the method is reliable. The regression equations are $y = 74.25x + 57.43$ ($R^2 = 0.999$), $y = 2076.66x + 233.88$ ($R^2 = 0.999$), and $y = 10632.71x + 337.38$ ($R^2 = 0.999$) for phenylalanine, tyrosine, and tryptophan respectively. The values of R^2 are all above 0.995, each R^2 deems a good linear fit for each analyte in its calibration range.

Predicted concentrations for the spiked validation sets using the PARAFAC-ALS method ($N = 6$) are given in Table 2. The root mean square errors of validation (RMSEVs) are 3.04, 0.07, and

0.02 $\mu\text{g mL}^{-1}$ respectively. The average spike recoveries (mean \pm standard deviation) are $98.5\pm 7.8\%$, $103.7\pm 6.9\%$, and $102.3\pm 7.9\%$ respectively. These results show that the proposed method is accurate and reliable enough for determining phenylalanine, tyrosine, and tryptophan in two different biofluids simultaneously.

For predicting the contents of phenylalanine in cell culture RPMI-1640, we prepared additional three prediction samples, each containing 5.00 ml cell culture in 10.00 mL. Table 3 gives the predicted contents of the three analytes in cell culture RPMI-1640 (which have been transformed by dilution factor) and analytical figures of merit. The normal levels of phenylalanine, tyrosine, and tryptophan in cell culture RPMI-1640 are 15.00, 20.00, and 5.00 $\mu\text{g mL}^{-1}$ respectively. From the prediction set 1, the predicted concentrations (mean \pm standard deviation) are 14.37 ± 0.37 , 21.27 ± 0.21 , and 4.96 ± 0.04 $\mu\text{g mL}^{-1}$ for phenylalanine, tyrosine, and tryptophan in cell culture RPMI-1640 respectively. Comparing with the real levels, the relative errors are -4.2% , 6.3% , and -0.8% respectively. The relative errors contain all errors in the whole method (from sampling to

prediction), which measures the error of the proposed method. For phenylalanine, tyrosine, and tryptophan, the SENs are 38.10, 255.87, and 239.79 $\text{mL } \mu\text{g}^{-1}$ respectively; the LODs are 0.50, 0.07, and 0.02 $\mu\text{g mL}^{-1}$ respectively; the LOQs are 1.51, 0.22, and 0.06 $\mu\text{g mL}^{-1}$ respectively.

The normal levels of phenylalanine, tyrosine, and tryptophan in human plasma are approximately the same and in the range 5 to 15 ppm⁵³, one can prepare about one-fifth or one-half (v: v) plasma sample to predict the real levels of phenylalanine, tyrosine, and tryptophan in human plasma. It is worth noting that the fluorescence intensities of spectral interferents (mostly are plasma protein) in about one-fifth or one-half (v: v) plasma sample will be very strong. However, the three-way calibration based EEMs used here has demonstrated that the direct determination of phenylalanine, tyrosine, and tryptophan in human plasma could be achieved, even in the presence of all uncalibrated interferents in plasma. Therefore, one can release the interferents in plasma to some extent by adding acetonitrile to plasma for protein precipitation.

Table 2 Predicted concentrations for the spiked validation sets using the PARAFAC-ALS method ($N = 6$).

Sample	Spiked concentrations ($\mu\text{g mL}^{-1}$)			Predicted concentrations ($\mu\text{g mL}^{-1}$) [Spike recovery %]		
	Phenylalanine	Tyrosine	Tryptophan	Phenylalanine	Tyrosine	Tryptophan
Validation set 1 (Cell culture)						
val01	6.00	0.40	0.10	6.32 [105.3]	0.45 [112.5]	0.12 [120.0]
val02	19.50	1.30	0.32	18.97 [97.3]	1.31 [100.8]	0.33 [103.1]
val03	33.00	2.20	0.55	31.23 [94.6]	2.18 [99.1]	0.54 [98.2]
val04	46.50	3.10	0.77	42.00 [90.3]	3.01 [97.1]	0.74 [96.1]
val05	60.00	4.00	1.00	53.33 [88.9]	3.92 [98.0]	0.95 [95.0]
Validation set 2 (Human plasma)						
val06	6.00	0.40	0.10	6.81 [113.5]	0.47 [117.5]	0.11 [110.0]
val07	19.50	1.30	0.32	20.63 [105.8]	1.41 [108.5]	0.34 [106.3]
val08	33.00	2.20	0.55	32.95 [99.8]	2.27 [103.2]	0.55 [100.0]
val09	46.50	3.10	0.77	44.91 [96.6]	3.15 [101.6]	0.76 [98.7]
val10	60.00	4.00	1.00	55.54 [92.6]	3.95 [98.8]	0.96 [96.0]
RMSEV ($\mu\text{g mL}^{-1}$)				3.04	0.07	0.02
Average spike recovery (mean \pm standard deviation)				$98.5\pm 7.8\%$	$103.7\pm 6.9\%$	$102.3\pm 7.9\%$

Table 3 Predicted contents of analytes in cell culture RPMI-1640 and analytical figures of merit using the PARAFAC-ALS method ($N = 6$).

Sample	Real concentrations ($\mu\text{g mL}^{-1}$)			Predicted concentrations ($\mu\text{g mL}^{-1}$)		
	Phenylalanine	Tyrosine	Tryptophan	Phenylalanine	Tyrosine	Tryptophan
Prediction set 1 (Cell culture RPMI-1640)						
Real	15.00	20.00	5.00			
pre01				14.79	21.06	4.92
pre02				14.24	21.28	4.95
pre03				14.08	21.47	5.00
RMSEP ($\mu\text{g mL}^{-1}$)				0.70	1.28	0.05
Content (mean \pm standard deviation) ($\mu\text{g mL}^{-1}$)				14.37 ± 0.37	21.27 ± 0.21	4.96 ± 0.04
Relative error				-4.2%	6.3%	-0.8%
SEN ($\text{mL } \mu\text{g}^{-1}$)				38.10	255.87	239.79
LOD ($\mu\text{g mL}^{-1}$)				0.50	0.07	0.02
LOQ ($\mu\text{g mL}^{-1}$)				1.51	0.22	0.06

4.5 Intraday and interday precision of the proposed method

The results of the intraday and interday precision for quantitative analysis of the analytes in two different systems simultaneously are presented in Table 4. Before calculating the RSD on average spike recovery of each experiment, the F -test was used to test the

precision of each spike experiment. All the calculated F values are less than critical value at the 98% confidence level, which means that there is no significant difference amongst the precisions of the three spike experiments. The interday precisions of the assay, calculated as the RSD on average spike recoveries from three repeated experiments performed in one given day, are 1.9%, 1.9%, and 1.2% for phenylalanine, tyrosine, and

tryptophan respectively. Likewise, the interday precisions are 0.8%, 3.3%, and 4.3% for phenylalanine, tyrosine, and tryptophan respectively.

All these results indicate that the proposed method can provide accurate and precise resolutions and predictions, with

"mathematical separation"⁵⁴ instead of chemical or physical separation, for the simultaneous determination of phenylalanine, tyrosine, and tryptophan in different sophisticated biofluids (cell culture and human plasma), even in the presence of three unknown, uncalibrated interferents.

Table 4 Intraday and interday precision of the assay for determining the analytes in multiple systems simultaneously.

	Average spike recovery					
	Intraday			Interday		
	Phenylalanine	Tyrosine	Tryptophan	Phenylalanine	Tyrosine	Tryptophan
Day3_1	98.5%	103.6%	101.5%	Day1	96.9%	103.8%
Day3_2	97.3%	107.7%	103.9%	Day2	97.9%	95.4%
Day3_3	96.4%	105.6%	102.8%	Day3_1	98.5%	101.5%
Mean	97.4%	105.6%	102.7%		97.8%	101.8%
RSD	1.9%	1.9%	1.2%		0.8%	3.3%
						4.3%

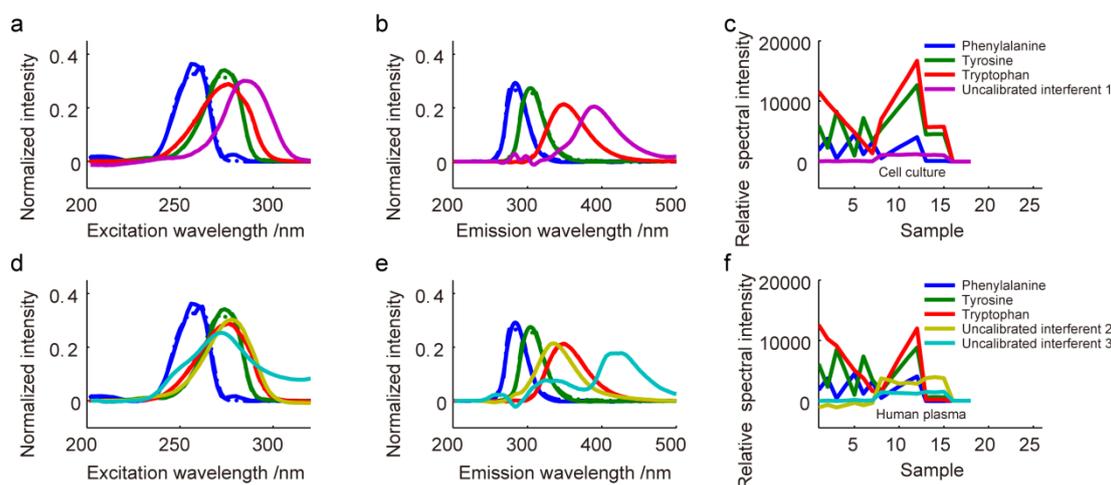


Fig. 6 Actual profiles (dash-dot lines) and decomposed profiles (solid lines) in each mode using the PARAFAC-ALS algorithm, individually predicting in cell culture ($N = 4$) and human plasma ($N = 5$) respectively.

4.6 The comparison between simultaneously predicting in multiple systems and individually predicting in single system

The "second-order advantage" is of great importance, since it allows analyte to be analyzed quantitatively even in the presence of uncalibrated interferent, which overcomes the prerequisite of one- and two-way calibration that calibration standards must be selected that are representative of the samples containing any interfering species. Direct analysis of analyte in complex system containing uncalibrated interferent has been fully discussed in theoretical and practical works, but simultaneous determination of analyte in different systems containing more than one uncalibrated interferents are rarely explored. By comparing the results of simultaneously predicting in cell culture and human plasma with that of individually predicting in one of them, we try to discuss the potentiality of three-way calibration method for determining analytes of interest in different systems simultaneously, to further explore the second-order advantages.

Figure 6 gives the actual profiles and decomposed profiles in

each mode using the PARAFAC-ALS algorithm, individually predicting in cell culture ($N = 4$) and human plasma ($N = 5$). Just like predicting in cell culture and human plasma simultaneously, the profiles of analytes in each mode were decomposed correctly and are consistent with the former ones, so do these of interferents.

Table 5 gives the results of comparison between predicting in cell culture and human plasma simultaneously and predicting in single system individually, by the paired t -test. For all the three analytes, both in cell culture and plasma, all the values of $t_{\text{Calculated}}$ are less than $t_{\text{Table}} = 3.75$ at the 98% confidence level. These paired t -test results indicate that, the predicted results between predicting in both systems simultaneously and predicting in single system individually have no significant difference, at the 98% confidence level. The satisfactory result means that, one could use a three-way calibration model to predict analytes of interest in multiple systems simultaneously, even in the presence of varying spectral interferents and backgrounds in different systems.

Cite this: DOI: 10.1039/c0xx00000x

www.rsc.org/xxxxxx

ARTICLE TYPE

Table 5 The comparison between predicting in multiple systems simultaneously and predicting in single system individually by the paired *t*-test.

Sample	Spiked concentrations ($\mu\text{g mL}^{-1}$)			Difference between predictions ($\mu\text{g mL}^{-1}$) Simultaneous result – Individual result		
	Phenylalanine	Tyrosine	Tryptophan	Phenylalanine	Tyrosine	Tryptophan
Validation set 1 (Cell culture)						
val01	6.00	0.40	0.10	-0.014	-0.003	-0.003
val02	19.50	1.30	0.32	0.030	-0.007	0.002
val03	33.00	2.20	0.55	0.029	-0.019	-0.003
val04	46.50	3.10	0.77	0.085	-0.051	-0.011
val05	60.00	4.00	1.00	0.120	-0.078	-0.019
^a $t_{\text{Calculated}}$				2.24	2.24	1.96
Validation set 2 (Human plasma)						
val06	6.00	0.40	0.10	0.033	0.001	-0.002
val07	19.50	1.30	0.32	0.095	0.005	-0.009
val08	33.00	2.20	0.55	0.167	0.006	-0.011
val09	46.50	3.10	0.77	0.224	0.011	-0.020
val10	60.00	4.00	1.00	0.290	0.014	-0.026
$t_{\text{Calculated}}$				2.24	3.13	3.48

^a $t_{\text{Calculated}}$ is computed by $t = \frac{|\bar{x}_d| \sqrt{n}}{s_d}$, where \bar{x}_d and s_d are the mean and standard deviation of these differences, respectively. At the 98% confidence level, $t_{\text{Table}} = 3.75$.

5. Conclusions

Three-way calibration method based on EEM fluorescence was applied for the direct determination of L-phenylalanine, L-tyrosine, and L-tryptophan in cell culture and human plasma simultaneously, despite the uncalibrated serious interferents in different systems. The proposed analytical assay has the advantages of being low cost, rapid, and environmentally friendly, with little chemical or physical separation. This is possible thanks to the second-order advantage which allows for quantitative analysis of analytes of interest in very complex systems not only containing unknown, uncalibrated interferents but also existing high collinear among spectra.

In addition, we discussed the potentiality of three-way calibration method for determining analytes of interest in different systems simultaneously, to further explore the second-order advantages. The satisfactory result means that, one could use a three-way calibration model to predict analytes of interest in multiple systems simultaneously, even in the presence of varying spectral interferents and backgrounds in different systems. These results obtained in this work indicate that the use of three-way calibration method coupled with EEMs array is a promising tool for multi-component simultaneous determination in multiple complex systems containing uncalibrated spectral interferents.

Acknowledgements

The authors gratefully acknowledge the National Natural Science Foundation of China (No. J1103312, No. 21175041), the National Basic Research Program (No. 2012CB910602), and the Foundation for Innovative Research Groups of NSFC (No. 21221003).

Notes and references

- ^a State Key Laboratory of Chemo/Biosensing and Chemometrics, College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, China. Fax: + 86 731 88821818; Tel: + 86 731 88821818; E-mail: hlwu@hnu.edu.cn
- A. Lehninger, D. L. Nelson and M. M. Cox, *Lehninger's Principles of Biochemistry*, 4th edn., WH Freeman and Company, New York, USA, 2005.
- J. Le Boucher, C. Charret, C. Coudray-Lucas, J. Giboudeau and L. Cynober, *Clin. Chem.*, 1997, **43**, 1421-1428.
- T. Soga and D. N. Heiger, *Anal. Chem.*, 2000, **72**, 1236-1241.
- D. M. Bailey, A. Hennig, V. D. Uzunova and W. M. Nau, *Chem-Eur. J.*, 2008, **14**, 6069-6077.
- F. Y. Tan, C. Tan, A. P. Zhao and M. L. Li, *J. Agr. Food. Chem.*, 2011, **59**, 10839-10847.
- J. R. Lakowicz, *Principles of fluorescence spectroscopy*, 3rd edn., Springer, New York, USA, 2006.
- S. Kudryasheva, *Anal. Bioanal. Chem.*, 2013, **405**, 3351-3358.
- C. Kang, H. L. Wu, Y. J. Yu, Y. J. Liu, S. R. Zhang, X. H. Zhang and R. Q. Yu, *Anal. Chim. Acta.*, 2013, **758**, 45-57.
- G. R. Li and L. J. Han, *Anal. Methods*, 2014, **6**, 426-432.
- H. X. Kuang, *Anal. Methods*, 2014.
- A. C. Olivieri, *Anal. Chem.*, 2008, **80**, 5713-5720.
- S. D. Brown, R. Tauler and B. Walczak, *Comprehensive Chemometrics: Chemical and Biochemical Data Analysis*, Elsevier, Amsterdam, the Netherlands, 2009.
- H. Wold and E. Lyttkens, *Bull. Intern. Statist. Inst. Proc.*, 37th London, 1969, **43**, 29-51.

- 1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
14. Y. Z. Liang, O. M. Kvalheim, H. R. Keller, D. L. Massart, P. Kiechle and F. Erni, *Anal. Chem.*, 1992, **64**, 946-953.
15. R. Tauler, B. Kowalski and S. Fleming, *Anal. Chem.*, 1993, **65**, 2040-2047.
16. X. G. Shao, G. R. Du, M. Jing and W. S. Cai, *Chemom. Intell. Lab. Syst.*, 2012, **114**, 44-49.
17. J. S. Luo, L. Z. Yu, Y. Z. Guo and M. L. Li, *Chemom. Intell. Lab. Syst.*, 2012, **110**, 163-167.
18. P. Gemperline, *Practical Guide to Chemometrics*, 2nd edn., CRC press, Boca Raton, FL, USA, 2006.
19. X. J. Tang, J. M. Xiao, Y. Z. Li, Z. N. Wen, Z. Fang and M. L. Li, *Chemom. Intell. Lab. Syst.*, 2012, **118**, 317-323.
20. A. R. Muroski, K. S. Booksh and M. Myrick, *Anal. Chem.*, 1996, **68**, 3534-3538.
21. K. S. Booksh, A. R. Muroski and M. Myrick, *Anal. Chem.*, 1996, **68**, 3539-3544.
22. E. Sanchez and B. R. Kowalski, *Anal. Chem.*, 1986, **58**, 496-499.
23. B. E. Wilson, W. Lindberg and B. R. Kowalski, *J. Am. Chem. Soc.*, 1989, **111**, 3797-3804.
24. P. Geladi, *Chemom. Intell. Lab. Syst.*, 1989, **7**, 11-30.
25. E. Sanchez and B. R. Kowalski, *J. Chemom.*, 1990, **4**, 29-45.
26. H. L. Wu, M. Shibukawa and K. Oguma, *J. Chemom.*, 1998, **12**, 1-26.
27. Z. P. Chen, H. L. Wu, J. H. Jiang, Y. Li and R. Q. Yu, *Chemom. Intell. Lab. Syst.*, 2000, **52**, 75-86.
28. Y. X. Tan, J. H. Jiang, H. L. Wu, H. Cui and R. Q. Yu, *Anal. Chim. Acta*, 2000, **412**, 195-202.
29. R. Tauler, *Chemom. Intell. Lab. Syst.*, 1995, **30**, 133-146.
30. J. A. Arancibia, A. C. Olivieri, D. B. Gil, A. E. Mansilla, I. Durán-Merás and A. M. de la Peña, *Chemom. Intell. Lab. Syst.*, 2006, **80**, 77-86.
31. T. G. Kolda and B. W. Bader, *SIAM Rev.*, 2009, **51**, 455-500.
32. E. Correa, H. Sletta, D. I. Ellis, S. Hoel, H. Ertesvåg, T. E. Ellingsen, S. Valla and R. Goodacre, *Anal. Bioanal. Chem.*, 2012, **403**, 2591-2599.
33. A. K. Smilde, *Chemom. Intell. Lab. Syst.*, 1992, **15**, 143-157.
34. K. S. Booksh and B. R. Kowalski, *Anal. Chem.*, 1994, **66**, 782-791.
35. G. M. Escandar, A. C. Olivieri, N. K. M. Faber, H. C. Goicoechea, A. Muñoz de la Peña and R. J. Poppi, *TrAC, Trends Anal. Chem.*, 2007, **26**, 752-765.
36. R. A. Harshman, *UCLA Working Papers in Phonetics*, 1970, **16**, 1-84.
37. J. D. Carroll and J.-J. Chang, *Psychometrika*, 1970, **35**, 283-319.
38. J. J. Liao, X. Ma, Y. D. Wen, Y. Wang, W. S. Cai and X. G. Shao, *Ind. Eng. Chem. Res.*, 2011, **50**, 7677-7681.
39. Y. N. Ni, Y. Gu and S. Kokot, *Chemom. Intell. Lab. Syst.*, 2012, **112**, 55-64.
40. A. C. Olivieri, *Anal. Methods*, 2012, **4**, 1876-1886.
41. H. Chen and J. E. Kenny, *Analyst*, 2012, **137**, 153-162.
42. K. R. Murphy, C. A. Stedmon, D. Graeber and R. Bro, *Anal. Methods*, 2013, **5**, 6557-6566.
43. F. Rabbani, H. Abdollahi and M. R. Hormozi-Nezhad, *Anal. Methods*, 2014.
44. J. B. Kruskal, *Linear Algebra Appl.*, 1977, **18**, 95-138.
45. A. C. Olivieri and N. K. M. Faber, *J. Chemom.*, 2005, **19**, 583-592.
46. A. C. Olivieri and N. K. M. Faber, *Chemom. Intell. Lab. Syst.*, 2004, **70**, 75-82.
47. D. C. Harris, *Quantitative chemical analysis*, 7th edn., W. H. Freeman and Company, New York, USA, 2007.
48. K. T. Fang, *Sci. Press, Beijing, China*, 1994.
49. K. T. Fang, D. K. Lin, P. Winker and Y. Zhang, *Technometrics*, 2000, **42**, 237-248.
50. M. Bahram, R. Bro, C. Stedmon and A. Afkhami, *J. Chemom.*, 2006, **20**, 99-105.
51. P. J. Gemperline, *J. Chem. Inf. Comp. Sci.*, 1984, **24**, 206-212.
52. R. Bro and H. A. Kiers, *J. Chemom.*, 2003, **17**, 274-286.
53. R. J. Wurtman, C. M. Rose, C. Chou and F. F. Larin, *New. Engl. J. Med.*, 1968, **279**, 171-175.
54. H. L. Wu, J. F. Nie, Y. J. Yu and R. Q. Yu, *Anal. Chim. Acta.*, 2009, **650**, 131-142.