Scientific Report
regarding the implementation of the project during the period December 16th 2015 – November 4th 2016

Project:	PN-II-ID-PCE-2011-3-0152
Project title:	Injection of large volumes of immiscible solvents with the mobile phase in high pressure liquid chromatography.
Contract number:	310/05.10.2011
Project Director:	Prof. Dr. Andrei Medvedovici
Time period for the report:	16.12.2015 – 04.11.2016

Due to the rescheduling of the initial project for an extended period (until November 4th 2016), this report covers the objectives and activities recorded in the project plan, Annex IV to the Grant Contract no. 310 / 05.10.2011.

The following objectives and activities are in discussion:

O3. Characterization of the hydrophobicity of solutes based on their ability to perform LVI:

O3A1. Series of immiscible diluents for progressive log P scales of target analytes.

O4. On-line automation issues in LVI of immiscible diluents for LC.

O4A2. SPE/LVI/LC issues.

O7. Dissemination of the results:

O7A1. Publication of experimental data in ISI journals.

O7A2. Participation in scientific conferences.

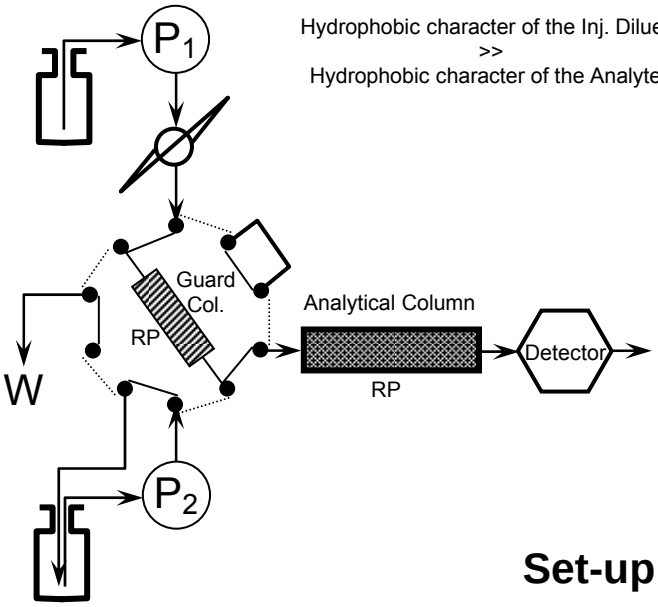
For O4A2:

Due to the fact that the previous publication in the project P. Lazăr, Ş. Udrescu, F. Tache, F. Albu, N. Grinberg, A. Medvedovici, *Revisiting large volume injection in non-miscible diluents: an on-line reversed phase supported liquid extraction / liquid chromatography scenario*, ***Analytical Methods*** 7 (2015) 342-352, undoubtedly clarify the intimate mechanism of LVI in mobile phase non-immiscible diluents and objectively demonstrated that the process is based on an on-line SLE/RPLC approach, the most straightforward way to realize automation is to split the analytical column in two separate ones, the first receiving the sample (including the diluent) and the latter selectively receiving only the analytes isolated in the SLE process.

The experimental set-ups were realized, as depicted in the following figures. All set-ups are based on a 10 ports high pressure rotating valve separating an SLE column from the analytical column it-self. Two high pressure pumps should be used, in order to perform experiments.

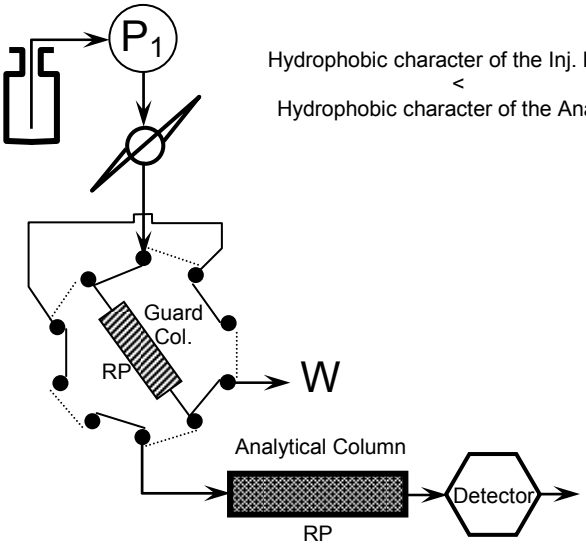
The set-up A corresponds to the situation in which the diluent (non-miscible with the mobile phase) has a much higher hydrophobic character compared to the analytes of interest. The first column (small length, i.e. 5 cm), filled with a RP stationary phase, acts a SLE cartridge. The second column (with a greater length, i.e. 15 cm) acts as an analytical column and is exploited in the same RP separation mechanism. The sample is loaded to the first column, with a water rich mobile phase and the valve in the position corresponding to the continuous lines in the figure. Analytes are first distributed in the diluent strongly retained on the stationary phase surface, next redistributed in the mobile phase (a typical SLE process) and transported and focussed in the analytical column head. Once the transport is realized, the valve switches in the dotted lines position. Pump 2 will be programmed to deliver 100% organic solvent, while Pump 1 will continue the initial composition profile. Analytes are thus separated in the chromatographic column, while the diluent is simultaneously back flushed from the SLE cartridge to waste.

The set-up B is based on the same principle, but it is designed for the case when the hydrophobic character of the diluent is less important compared to analytes.



Hydrophobic character of the Inj. Diluent
>>
Hydrophobic character of the Analytes

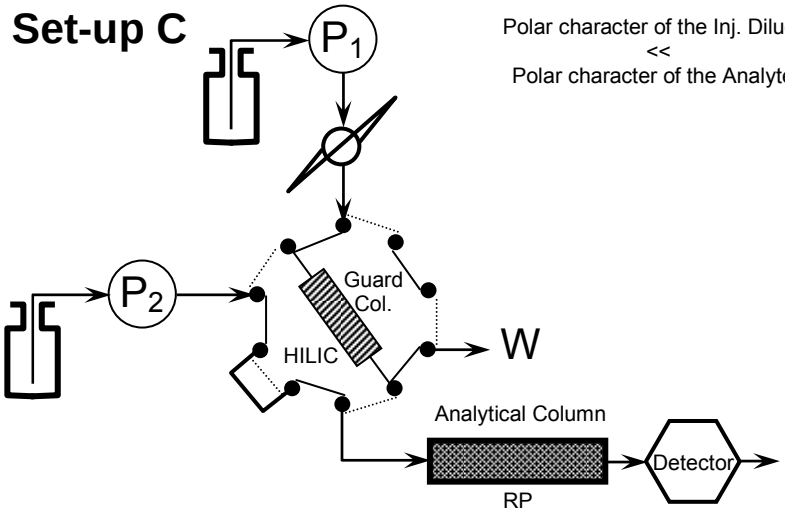
Set-up A



Hydrophobic character of the Inj. Diluent
<
Hydrophobic character of the Analytes

Set-up B

Set-up C



Polar character of the Inj. Diluent
<<
Polar character of the Analytes

This time a single high pressure HPLC pump is needed. The sample is loaded to the first column, which is acting as an SPE cartridge. The analytes are strongly retained in the stationary phase, while the diluent is directed to waste. After the diluent removal, the valve switches to the dotted lines position. The flow is redirected to the analytical column, which focus the analytes from the SPE column and separate them.

Finally, the set-up C corresponds to the case in which the polar character of the diluent is less important compared to the analytes. This time, the first column (acting as a SPE cartridge), is exploited in the HILIC mode, while the analytical column is exploited in the RP mode. In the first stage, when sample is loaded, with an organic solvent rich mobile phase composition and the valve is in the position indicated by the continuous line, the diluent is diverted to waste (the diluent, being less polar, the analytes elutes first from the HILIC column). Meanwhile, pump 2 feeds the analytical column with a typical RP mobile phase (water rich one). After diluent removal, the valve switches (position corresponding to the dotted lines). The analytes are desorbed by the water rich mobile phase and focused in the head of the analytical column, where the chromatographic separation is achieved.

The major draw-back of the presented experimental set-ups consist in a laborious development phase. The valve switching time is injection volume dependent, and should be attentively optimized. The composition of the mobile phases delivered by the two pumps or the gradient profile of pump 1 in set-up B should be also attentively optimized in order to achieve focusing of the analytes in the head of the analytical column.

Such approaches may be usefully developed only for routine applications necessitating the analysis of a large number of samples.

For O3A1:

Due to the fact that this subject was extensively discussed in the **Report on 2015**, we preferred to continue the work on the new chemometric algorithm developed by us, namely the Linear Regression Algorithm. For such a purpose, we found an application corresponding to **O4A6** in the planning of the present project. This topic will be further discussed in the present report. Experimental data will be sent for publication until the end of the 2016 year.

The discrimination power of a novel algorithm based on linear regression (LRA) for evaluation of the green teas γ-irradiation effects was assessed. The evaluation was holistically applied directly to (+/-) ESI/MS, and RPLC/UV chromatograms, without involving any structural attribution and/or assay of the existing components. Five types of green teas (irradiation doses of 0, 10 and 25 kGy) were considered. Extraction in ethanol and heated water were used. To ensure an increased definition of the profiles being compared, the LRA approach was applied on pairs of large experimental data series resulting from high frequency/high resolution acquisition rates, the resulting slopes, intercepts and correlation coefficients being considered as variables retaining the information contained in the raw data. The information contained by the input data varied as following: (+)ESI/MS spectra > (-)ESI/MS spectra > RPLC/UV chromatograms.

Five commercial brands of green teas were obtained from the local market (Shutao, Twinnings, Lipton, Celmar, Carrefour). Products were labeled with a disclosure referring to any γ-irradiation treatments during production. These samples are further referred in the text as numbers, from 1 to 5.

A GC-5000 Co-60 gamma irradiator was used for irradiation. Irradiation was carried out in the presence of air, at room temperature. An ECB dosimetric system was used for dose values of samples, which are expressed as absorbed dose in water. The green tea samples were irradiated at the doses of (10.2 ± 0.9) kGy and (25.5 ± 2.3) kGy, respectively. The irradiation time was 1h 40 min and 4h 10 min, respectively.

Water infused samples were obtained through addition of 2 mL of boiling water to 100 mg of solid material (no mechanical processes were used to reduce the sample size or to control particle size). Vials previously tightly closed, are kept at 80 °C for 1 hour. After cooling and centrifugation at 6000×g, the supernatant was isolated, placed in an injection vial, sealed and used for experiments.

Ethanol extracts were obtained through addition of 2 mL of ethanol to 100 mg of solid sample (without any previous preparation of the solid material), at room temperature. The sealed vials were vortexed for 1 min at 2500 rpm and kept in dark at room temperature for 7 days. After a final vortexing

period of 5 min at 2500 rpm and centrifugation at 6000×g, the supernatant was isolated in an injection vial, tightly sealed.

For the RPLC separations, an Eclipse XDB-C18 column, 150 mm length × 4.6 mm internal diameter × 3.5 μm particle size (Agilent Technologies) was used. The mobile phase was composed from acetonitrile (solvent A) and aqueous 0.1% formic acid (v/v) (solvent B) and had a flow rate of 1 mL min^{-1}. The column was thermostated at 25 °C. The injected volume of the sample was 10 μL. The analytical wavelength was fixed at 220 nm (± 2 nm spectral window), with a reference spectral window of 480 ± 5 nm. The auto-balancing option was automatically set to the detector before each injection (in order to fix the 0 point at the starting of the chromatogram acquisition). The following gradient profiles were used (including the re-equilibration step): a) for aqueous infused samples; min/solvent A % (v/v): 0.0/2%; 5.0/2%; 10.0/12%; 15.0/32%; 20.0/72%; 20.01/100%; 24.0/100%; 24.01/2%; 34.0/2%; b) for extracts in ethanol; min/solvent A % (v/v): 0.0/5%; 5.0/10%; 20.0/35%; 25.0/55%; 30.0/100%; 35.0/100%; 35.01/5%; 45.0/5%. Consequently, the duration of the chromatograms resulting from aqueous infused samples was 23 min and 35 min for ethanol extracts, respectively. The data acquisition frequency of the detector was set at 5 Hz, meaning one measurement at every 0.198 s (or 0.0033 min, respectively). Consequently, a chromatogram of an aqueous infused sample is retained as a 6900 values length series, while chromatograms of the ethanol extracts have 10500 values for each series.

A direct infusion of the samples in the electrospray ion source operated in either positive or negative ionization mode was used for generating mass spectra. A mixture of acetonitrile and 0.1% aqueous formic acid solution (v/v) in a volumetric ratio 1/1, at a flow rate of 0.5 mL min^{-1} was used as carrier flow. The autosampler was connected to the ion source inlet through a stainless steel capillary, 2 m length and 0.12 mm internal diameter, maintained in the column thermostat at 35 °C. The sample injected volume was 10 μL. The operational parameters of the electrospray ion source were: drying gas temperature - 350 °C; drying gas flow - 8 L/min; pressure of the nebulising gas - 40 psi; potential applied to the extraction capillary - (+) or (-) 4000 V. The operating parameters of the mass analyzers were: scan type - MS2 scan; ion polarity - positive or negative; data storage - profile; peak width - 0.07; declustering potential (fragmentor) - 135 V; scanning interval - m/z from 20 to 1100 a.m.u.; collision cell voltage - 7 V; dwell time - 500 ms; electron multiplier voltage - 400 V. The spectral acquisition resolution was 0.1 a.m.u. The mass spectrum characterizing each sample represented an average of spectra acquired in the interval 0.2 and 0.6 min after infusion of the bulk in the ion source. Each injection takes 6 min, in order to eliminate the carryover effect from the ion source. Each mass spectra resulting from this experimental approach was digitized in series of 10800 values.

Five numerical data series are obtained for each type of tea and degree of irradiation. Exception is made by RPLC/UV analysis of aqueous infused and ethanol extracts corresponding to 10 and 25 kGy doses (in these cases a single data series is available).

For the five numerical series corresponding to a tea sample and degree of irradiation, one can reciprocally compare the data through linear regression analysis. Each linear correlation is characterized by a slope, an intercept and a correlation coefficient. For the 20 possible arrangements (arrangement of 5 objects taken as couples) it is then possible to calculate the mean values of the linear regression parameters (slope, intercept and correlation coefficient) as well as their corresponding standard deviations (s_D).

When comparing two sets of five numerical series corresponding to different irradiation doses, the sets corresponding to the lowest irradiation dose were kept on the abscissa. A set of 25 linear regression parameters (slopes, intercepts and correlation coefficients) are thus created, then averaged and the standard deviations being calculated.

Consequently, one can determine an ellipsoidal volume in the three dimensional Cartesian space (slopes placed on the *Ox* axis, intercepts on the *Oy* axis, and correlation coefficients on the *Oz* axis), for each analyzed sample or comparison of samples, having as the centre the mean values of the linear regression parameters and as semi-diameters *a, b, c* the segments corresponding to two times the standard deviations ($2\times s_D$) on each direction of the Cartesian space.

Figure 1 illustrates the above statements, in the case of tea 1 (non-irradiated and irradiated with 10 and 25 kGy doses), when analyzing ethanol extracts through (+) ESI/MS. Four ellipsoidal

volumes were thus created, one corresponding to tea 1 non-irradiated (1_0 kGy), two others corresponding to comparisons of tea 1 non-irradiated and tea 1 irradiated at 10 and 25 kGy doses, respectively (1_0 vs 10 kGy and 1_0 vs 25 kGy) and one corresponding to comparison of tea 1 irradiated at 10 and 25 kGy doses (1_10 vs 25 kGy).

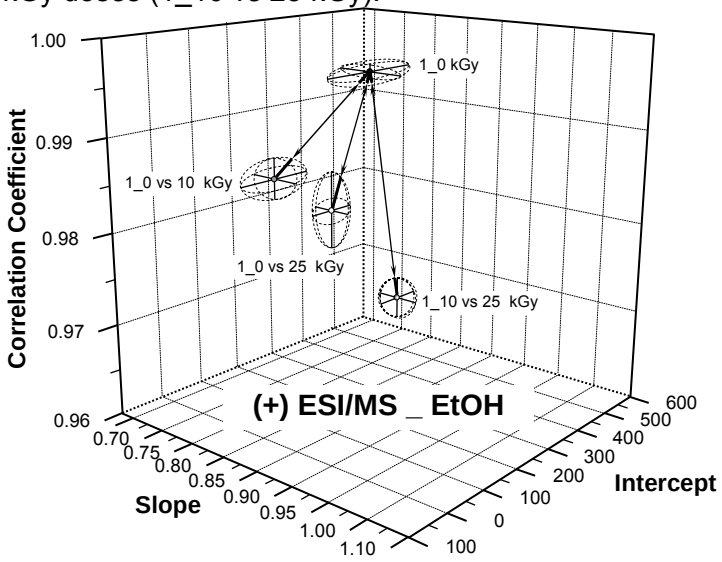


Figure 1. Results obtained through application of the Linear Regression Algorithm (LRA) to (+) ESI/MS spectra of the ethanol extracts from non-irradiated and irradiated (10 and 25 kGy doses) green tea 1, represented in the Cartesian 3D space.

The true distance between two ellipsoids (*d*) is given by the distance *D* between their centers from which we need to subtract the radii of the two ellipsoids on the direction of the linkage between their centers. The mathematical relationships required for computation of such true distances are given in Figure 2**.**

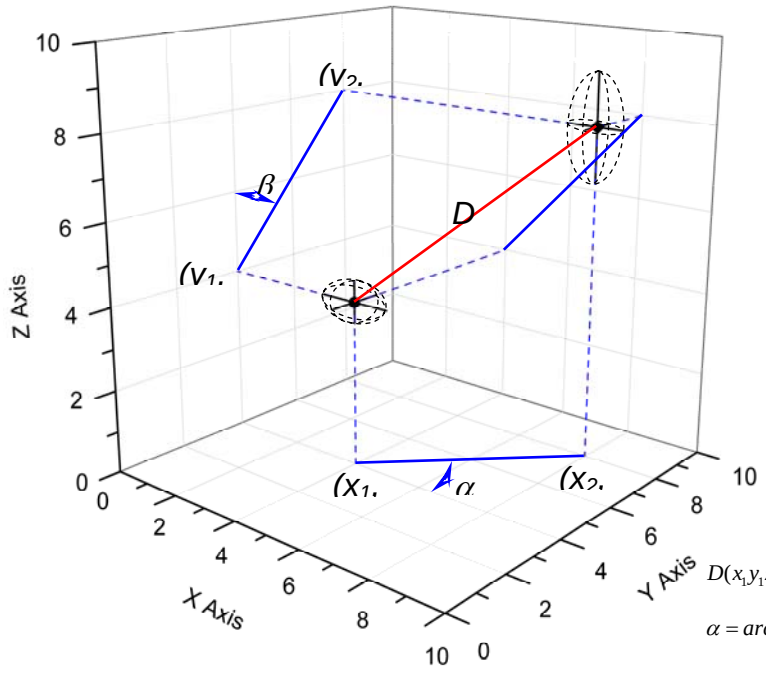
If two ellipsoids are clearly separated, the true distance between them, *d* should be higher than 0. A negative true distance between ellipsoids means that they are overlapping. Higher is the true distance between two ellipsoidal volumes, higher is the difference between the compared samples.

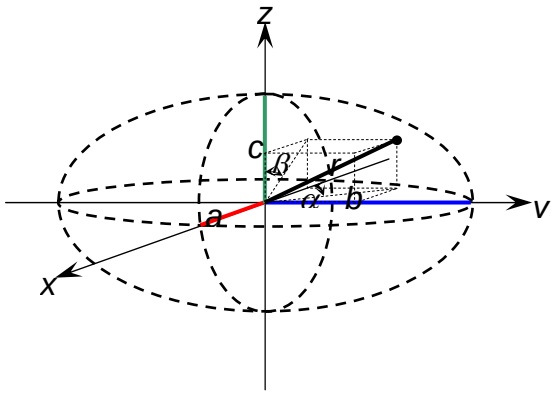
As particular cases, when comparing RPLC/UV chromatograms of teas non-irradiated and irradiated at 10 or 25 kGy doses, mean linear regression parameters and their corresponding normal variation intervals are provided by a set of only 5 discrete linear regressions. While comparing RPLC/UV chromatograms of teas irradiated at 10 and 25 kGy, respectively, only one regression is obtained. Consequently, the distance is measured between 2 points in the 3D space.

If a volume illustrating the comparison between two different irradiation states overlays the volume generated through comparison of the input data series acquired for a single irradiated state (for instance the non-irradiated one), means that the irradiation has no significant influence on the sample (compared to the instrumental variability produced during data acquisition).

The graphical representation of the LRA results (see Figure 1) is somehow difficult to achieve, in the absence of specialized 3D graphical softwares. The 3D representations may be difficult to be optimally visualized in the absence of graphical rotational features. As illustrated in Figure 2 for the (-) ESI/MS spectra of the aqueous extracts of non-irradiated and irradiated tea 3 samples, the projections in the three planes of the Cartesian space may successfully replace the 3D representations.

The computation of the true distances between the ellipsoidal volumes determined through the LRA approach allows a more objective interpretation of the experimental results. The results obtained through application of the LRA approach to the experimental data sets are also enlisted in Table 1.





$x_i = \overline{B_i}$; B_i = slopes

$y_i = \overline{A_i}$; A_i = intereceptѕ

$z_i = \overline{C_i}$; C_i = correlation coefficients

$a = 2 \times s_B$; $b = 2 \times s_A$; $c = 2 \times s_C$; s = standard deviation

$$D(x_1y_1z_1, x_2y_2z_2) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2 + (z_2 - z_1)^2}$$

$$\alpha = arctg\left(\frac{y_2 - y_1}{x_2 - x_1}\right)$$

$$\beta = arctg\left(\frac{y_2 - y_1}{z_2 - z_1}\right)$$

$$r = \frac{a \times b \times c}{\sqrt{b^2 \times c^2 \times \cos^2\alpha \times \sin^2\beta + a^2 \times c^2 \sin^2\alpha \times \sin^2\beta + a^2 \times b^2 \times \cos^2\beta}}$$

$$d_{1,2} = d(x_1y_1z_1, x_2y_2z_2) - r_1 - r_2$$

From Table 1, one can conclude that (+) ESI/MS spectra of the ethanolic extracts, as well as (-) ESI/MS spectra of both aqueous and ethanol extracts, allows discrimination between the irradiation states of the samples. The RPLC/UV chromatograms of the aqueous extracts fails to separate non-irradiated against 10 kGy irradiation dose for teas 1 and 5. The (+) ESI/MS spectra of the aqueous extracts fails to discriminate between the different irradiation states of tea 4 (0 against 10 kGy and 0 against 25 kGy, respectively). RPLC/UV chromatograms of the ethanolic extracts achieves discrimination according to the irradiation status with a single exception (0 against 10 kGy, for tea 3), although the two ellipsoids are nearly tangent (the calculated true distance is practically null).

General conclusions

During the period December 16th 2015 – November 4th 2016, the project team realized 1 ISI publication, as it follows:

1. E. Iorgulescu, V.A. Voicu, C. Sârbu, F. Tache, F. Albu, A. Medvedovici, Experimental variability and data pre-processing as factors affecting the discrimination power of some chemometric approaches(PCA, CA and a new algorithm based on linear regression) applied to (±) ESI/MS and RPLC/UV data: Application on green tea extracts, ***Talanta*** 155 (2016) 133-144 (I.F. = 3,545).

The results presented in the above mentioned publications correspond to the following objectives (O) and activities (A) of the project: Publication 1. O6A4;

The dissemination of the results being obtained throughout the project plan completion was also achieved:

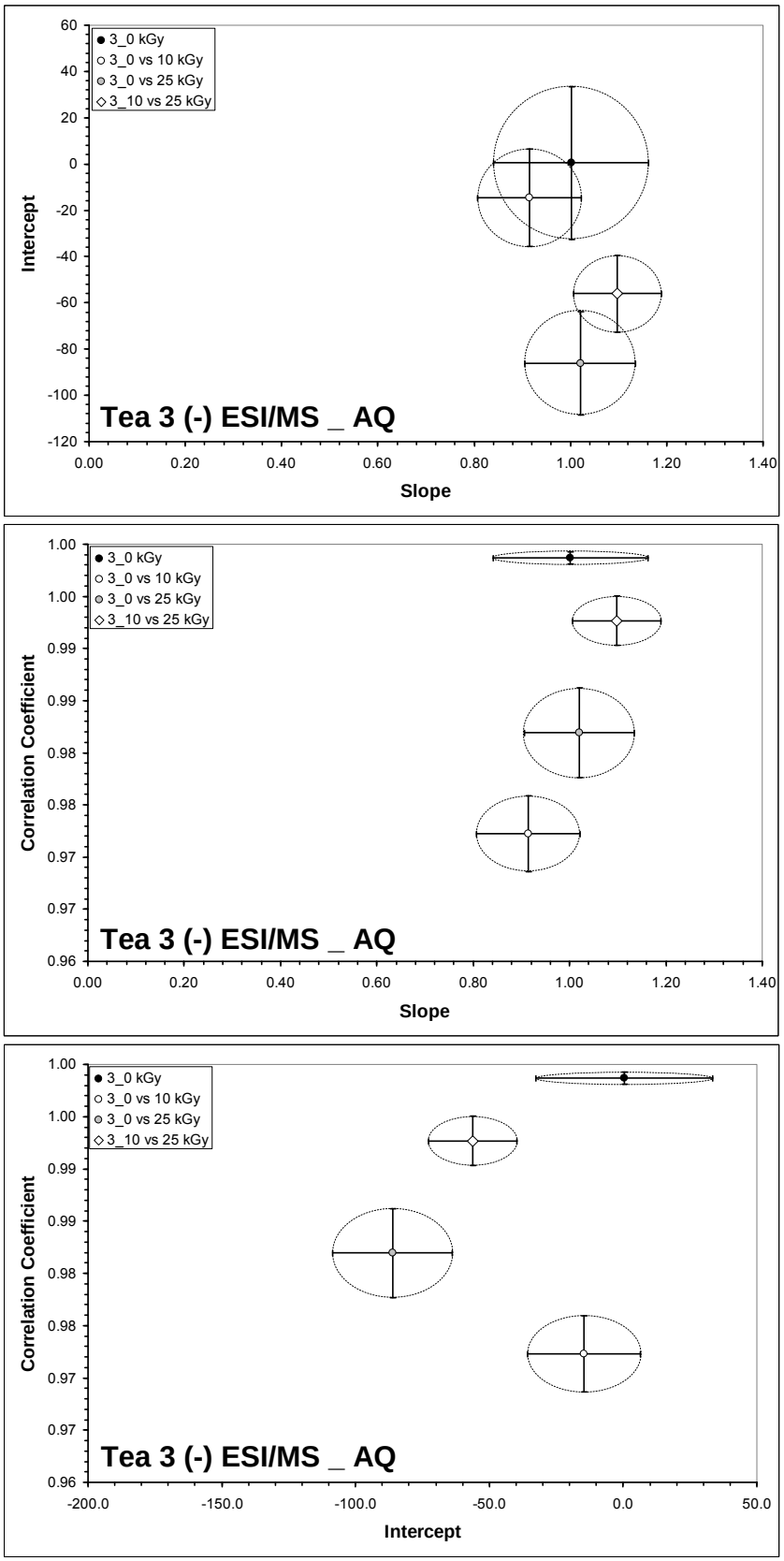


Figure 2. Projections in the three planes of the Cartesian space of the data resulting from the application of LRA to (-) ESI/MS spectra of the aqueous extracts of tea 3, non-irradiated or irradiated at 10 and 25 kGy doses.

Table 1. LRA ability to discriminate between non-irradiated and irradiated green teas at different doses (10 and 25 kGy), taking as input data the RPLC/UV chromatograms and (+)/(-) ESI/MS spectra of aqueous and ethanol extracts.

Experimental technique		RPLC/UV_AQ		RPLC/UV_EtOH		(+) ESI/MS_AQ		(+) ESI/MS_EtOH		(-) ESI/MS_AQ		(-) ESI/MS_EtOH	
1	1_0 vs 10 kGy	-0.30	*NS*	0.10	*S*	24.4	*S*	3.5	*S*	56.4	*CS*	42.3	*CS*
2	1_0 vs 25 kGy	0.20	*S*	0.50	*S*	10.3	*S*	417.2	*CS*	26.9	*S*	93.0	*CS*
3	1_10 vs 25 kGy	1.00	*CS*	0.60	*S*	20.3	*S*	474.4	*CS*	7.6	*S*	135.7	*CS*
4	2_ 0 vs 10 kGy	0.24	*S*	0.06	*S*	43.0	*S*	539.9	*CS*	15.7	*S*	79.8	*CS*
5	2_ 0 vs 25 kGy	1.09	*CS*	2.66	*CS*	575.7	*CS*	785.4	*CS*	9.4	*S*	35.7	*S*
6	2_ 10 vs 25 kGy	0.73	*CS*	2.87	*CS*	705.4	*CS*	427.0	*CS*	25.9	*S*	30.0	*S*
7	3_ 0 vs 10 kGy	0.10	*S*	-0.001	*NS*	266.0	*CS*	109.0	*S*	12.6	*S*	9.2	*S*
8	3_ 0 vs 25 kGy	1.70	*CS*	0.05	*S*	199.2	*S*	27.1	*S*	67.9	*CS*	11.6	*S*
9	3_ 10 vs 25 kGy	1.33	*CS*	0.03	*S*	5.2	*S*	59.3	*S*	38.3	*S*	0.1	*S*
10	4_ 0 vs 10 kGy	0.58	*CS*	0.11	*S*	-50.1	*NS*	10.3	*S*	115.6	*CS*	8.6	*S*
11	4_ 0 vs 25 kGy	0.26	*S*	0.22	*S*	-62.0	*NS*	61.9	*CS*	55.5	*CS*	1.4	*S*
12	4_ 10 vs 25 kGy	1.44	*CS*	0.12	*S*	87.1	*S*	87.1	*CS*	13.7	*S*	11.8	*S*
13	5_ 0 vs 10 kGy	-0.49	*NS*	0.13	*S*	14.7	*S*	51.6	*S*	163.5	*CS*	4.5	*S*
14	5_ 0 vs 25 kGy	0.15	*S*	0.17	*S*	17.9	*S*	305.1	*CS*	5.1	*S*	70.5	*CS*
15	5_ 10 vs 25 kGy	0.50	*S*	0.22	*S*	63.1	*S*	191.1	*S*	84.9	*CS*	98.5	*CS*

NS - Not Separated; S - Separated; CS - Clearly Separated

in Applied Chemistry, May 26th-28th, Constanta, Romania, 2016, Conference Program, pg. 7.

Symposium 1. ***New chemometric approach based on linear regression for holistic comparison of analytical data***; A. Medvedovici; International Conference Chimia 2016, New Trends

Symposium 2. ***Chemometric approaches for characterization of complex mixtures of natural origins used as active ingredients in dietary supplements***; A. Medvedovici, V. Voicu; International Conference "From Science to Guidance and Practice" June, 6th-7th, Bucharest, Romania, 2016, 2nd FSPG Program, Day 1, pg. 2.

The publication and conferences above mentioned are contributing to objectives O6A4 and O7A1of the project, referring to dissemination of the research results. Above mentioned publication include in the section *Acknowledgements* the indicative of the present grant. Power Point presentations used during oral presentations are also containing in the last slide the indicative of the present grant.

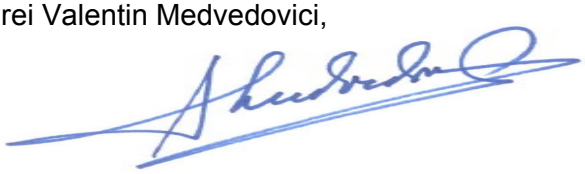
Publications (first page and the page with the Acknowledgements section) were posted on the project site http://www.chimie.unibuc.ro/cercetare/analitica /lvihplc/index.htm. Power Point presentations are also posted (in their entirety) on the same site.

Management and permanent up-dating of the project site responds equally to Objective 8.

The experimental data, the published manuscripts and plenary lectures contributing to the dissemination of results comply to objectives and activities redistributed according to project rescheduling plan (Annex IV of the financing contract nr. 310/2011). The Romanian version of the present document was up-loaded on the UEFISCDI site.

October 12th, 2016

Project Director,
Professor Dr. Andrei Valentin Medvedovici,



Address of the Project Director:
Department of Analytical Chemistry
Faculty of Chemistry
University of Bucharest
Panduri Ave. # 90-92, Bucharest-050663, Romania
E-mail: avmedved@yahoo.com
Fax: 021 4102279
Tel.: 072 3330423

Talanta 155 (2016) 133–144

Contents lists available at ScienceDirect

Talanta

journal homepage: www.elsevier.com/locate/talanta

Experimental variability and data pre-processing as factors affecting the discrimination power of some chemometric approaches (PCA, CA and a new algorithm based on linear regression) applied to (+/−)ESI/MS and RPLC/UV data: Application on green tea extracts

E. Iorgulescu [a], V.A. Voicu [b,c], C. Sârbu [d], F. Tache [a], F. Albu [e], A. Medvedovici [a,*]

[a] University of Bucharest, Faculty of Chemistry, Department of Analytical Chemistry, Panduri Ave., no. 90, Bucharest 050663, Romania
[b] Romanian Academy, Medical Science Section, Calea Victoriei no. 125, Bucharest 010071, Romania
[c] University of Medicine and Pharmacy "Carol Davila", Department of Pharmacology, Toxicology and Clinical Psychopharmacology, #8 Florescu St., Bucharest 014461, Romania
[d] Babeş-Bolyai University, Faculty of Chemistry and Chemical Engineering, Department of Chemistry, Arany Janos Street, no. 11, Cluj-Napoca 400028, Romania
[e] Analytical Application Laboratory, Agilrom, # 405 Th. Pallady Ave., Bucharest 032266, Romania

ARTICLE INFO

Article history:
Received 17 March 2016
Received in revised form
15 April 2016
Accepted 19 April 2016
Available online 21 April 2016

Keywords:
Multivariate methods
(+/−) ESI-MS
RPLC-UV
Influence of the experimental variability
Influence of data pre-processing
Linear regression analysis
Green teas

ABSTRACT

The influence of the experimental variability (instrumental repeatability, instrumental intermediate precision and sample preparation variability) and data pre-processing (normalization, peak alignment, background subtraction) on the discrimination power of multivariate data analysis methods (Principal Component Analysis -PCA- and Cluster Analysis -CA-) as well as a new algorithm based on linear regression was studied. Data used in the study were obtained through positive or negative ion monitoring electrospray mass spectrometry (+/−ESI/MS) and reversed phase liquid chromatography/UV spectrometric detection (RPLC/UV) applied to green tea extracts. Extractions in ethanol and heated water infusion were used as sample preparation procedures. The multivariate methods were directly applied to mass spectra and chromatograms, involving strictly a holistic comparison of shapes, without assignment of any structural identity to compounds. An alternative data interpretation based on linear regression analysis mutually applied to data series is also discussed. Slopes, intercepts and correlation coefficients produced by the linear regression analysis applied on pairs of very large experimental data series successfully retain information resulting from high frequency instrumental acquisition rates, obviously better defining the profiles being compared. Consequently, each type of sample or comparison between samples produces in the Cartesian space an ellipsoidal volume defined by the normal variation intervals of the slope, intercept and correlation coefficient. Distances between volumes graphically illustrates (dis) similarities between compared data. The instrumental intermediate precision had the major effect on the discrimination power of the multivariate data analysis methods. Mass spectra produced through ionization from liquid state in atmospheric pressure conditions of bulk complex mixtures resulting from extracted materials of natural origins provided an excellent data basis for multivariate analysis methods, equivalent to data resulting from chromatographic separations. The alternative evaluation of very large data series based on linear regression analysis produced information equivalent to results obtained through application of PCA an CA.

1. Introduction

Extracts from tea leaves (*Camellia sinensis*) contain a large variety of compounds belonging to various chemical classes [1,2]. Different analytical investigation techniques were used for characterization of such complex mixtures. This includes gas and liquid chromatography, as single or multi-dimensional approaches [3–6], coupled to various detection techniques for unravelling organic compounds patterns as well as atomic absorption or optical emission/mass spectrometry inductively coupled plasma techniques for determination of the multi-elemental profiles [7–9]. Electrochemistry [10], anti-oxidant activity and chelating capacity [11], olfactory properties [12,13], texture properties (through image processing) [14], high resolution melting assay [15], UV [16], NIR [17] and DART-MS [18] spectrometric

* Corresponding author.
E-mail address: avmedved@yahoo.com (A. Medvedovici).

http://dx.doi.org/10.1016/j.talanta.2016.04.042
0039-9140/

the cumulate effects of the sample preparation, intra and interday instrumental variability made impossible the distinction between the compared data. Background subtraction of the (+)ESI/MS spectra produced an increase of the measured true distances, while for the (−)ESI/MS spectra the differences are minor. However, one can conclude that the data pre-processing modes do not add any additional discrimination power to the proposed evaluation algorithm.

4. Conclusions

Mass spectra obtained after ionization from the liquid state, under atmospheric pressure working conditions (either with positive or negative ion monitoring) of infused bulk solutions of complex mixtures (resulting from extraction of natural origin materials) are confirmed to represent an excellent data basis for making discrimination among samples, when applying chemometric algorithms, such as PCA or CA. Separation of compounds via a chromatographic method using a "classical" detection approach (such as UV spectrometry) are equivalent to mass spectra obtained from the bulk mixture, from the point of view of the chemometric algorithms. Comparison of chromatograms and mass spectra can be made with respect to the experimental shapes only, without any structural attribution of the compounds in the mixtures (holistic approaches). Grouping of the samples according to different chemometric approaches may be influenced by the sample preparation algorithm or the analytical investigational technique.

Instrumental repeatability and sample preparation induced variability are affecting to a less extent the chemometric algorithms used for discrimination between samples, compared to the instrumental intermediate precision. Mass spectrometric data are more influenced by the instrumental intermediate precision than chromatographic data. Variability issues should be considered and included in the experimental design of the analytical stages, to better assess and validate the dissimilarities between compared samples. It is recommended that data acquisition should be made in a single experimental session. MS experiments have the advantage to be performed faster compared to LC approaches.

An alternative data processing algorithm for discriminating against the samples was proposed. This algorithm is based on the linear regression model applied mutually to experimental data series. Each sample or comparison between two samples is thus represented as an ellipsoidal volume in the 3D Cartesian space, having the slope, the intercept and the correlation coefficient plotted on the three coordinates. These volumes are determined by the normal variation intervals of the slopes, intercepts and correlation coefficients. In fact, these normal variation intervals are retrieving the information contained in very large data series (better reproducing the experimental profiles being acquired), compared to the limited ability of the chemometric PCA and CA approaches to accept data series higher than 1000 items. The proposed algorithm does not require a specialized statistic software.

Data pre-processing options, as normalization, peak alignment and baseline subtraction modes, are increasing the absolute distances in the principal components space, but not necessarily increase the discrimination power of the PCA and CA algorithms. For the LRA algorithm, normalization reduces the absolute measured distances, while background subtraction produced similar or enhanced distances. However, data pre-processing modes are not certainly improving the discrimination power of the LRA approach.

The new algorithm based on Linear Regression illustrates the way the holistic evaluations may contribute to characterization of very complex mixtures of natural origins. This may represent a straightforward issue for evaluation of the quality characteristics of complex mixtures, with a potential impact on the dietary supplements and functional food markets, representing a fast and relatively cheap way for the quality control & evaluation over production batches.

Acknowledgements

This work was possible with the financial support offered by the Romanian Ministry of Education and Research, through research grants PN-II-ID-PCE-2011–3-152 and PN-II-ID-PCE-2011-3-0366.

Appendix A. Supplementary material

Supplementary data associated with this article can be found in the online version at http://dx.doi.org/10.1016/j.talanta.2016.04.042.

References

[1] Tea in Health and Diseases Prevention, in: V.R. Preedy (Ed.), Elsevier Inc. Academic Press, London, 2013.
[2] J. Zhao, J.W. Deng, Y.W. Chen, S.P. Li, Advanced phytochemical analysis of herbal tea in China, J. Chromatogr. A 1313 (2013) 2–23.
[3] C.T. Scoparo, L.M. de Souza, N. Dartora, G.L. Sassaki, P.A.J. Gorin, M. Iacomini, Analysis of Camellia sinensis green and black teas via ultra high performance liquid chromatography assisted by liquid–liquid partition and two-dimensional liquid chromatography (size exclusion × reversed phase), J. Chromatogr. A 1222 (2012) 29–37.
[4] L. Zhang, Z. Zeng, C. Zhao, H. Kong, X. Lu, G. Xu, A comparative study of volatile components in green, oolong and black teas by using comprehensive two-dimensional gas chromatography–time-of-flight mass spectrometry and multivariate data analysis, J. Chromatogr. A 1313 (2013) 245–252.
[5] L.F. Wang, J.Y. Lee, J.O. Chung, J.H. Baik, S. So, S.K. Park, Discrimination of teas with different degrees of fermentation by SPME–GC analysis of the characteristic volatile flavour compounds, Food Chem. 109 (2008) 196–206.
[6] Y. Zhao, P. Chen, L. Lin, J.M. Harnly, L. Yu, Z. Li, Tentative identification, quantitation, and principal component analysis of green pu-erh, green, and white teas using UPLC/DAD/MS, Food Chem. 126 (2011) 1269–1277.
[7] A. Szymczycha-Madeja, M. Welna, P. Pohl, Determination of essential and non-essential elements in green and black teas by FAAS and ICP OES simplified – multivariate classification of different tea products, Microchem. J. 121 (2015) 122–129.
[8] A. Szymczycha-Madeja, M. Welna, P. Pohl, Elemental analysis of teas and their infusions by spectrometric methods, Trends Anal. Chem. 35 (2012) 165–181.
[9] P.H.G.D. Diniz, M.F. Pistonesi, M.B. Alvarez, B.S.F. Band, M.C.U. de Araujo, Simplified tea classification based on a reduced chemical composition profile via successive projections algorithm linear discriminant analysis (SPA-LDA), J. Food Comp. Anal. 39 (2015) 103–110.
[10] C. Bardpho, P. Rattamarat, W. Siangproh, O. Chailapakul, Ultra-high performance liquid chromatographic determination of antioxidants in teas using ink jet-printed graphene–polyaniline electrode, Talanta (2015), http://dx.doi.org/10.1016/j.talanta.2015.05.020.
[11] E. Damiani, T. Bacchetti, L. Padella, L. Tiano, P. Carloni, Antioxidant activity of different white teas: comparison of hot and cold tea infusions, J. Food Comp. Anal. 33 (2014) 59–66.
[12] B. Tudu, S. Ghosh, A.K. Bag, D. Ghosh, N. Bhattacharyya, R. Bandyopadhyay, Incremental FCM technique for black tea quality evaluation using an electronic nose, Fuz. Inf. Eng. 7 (2015) 275–289.
[13] V. Kraujalyte, E. Pelvan, C. Alasalvar, Volatile compounds and sensory characteristics of various instant teas produced from black tea, Food Chem. 194 (2016) 864–872.
[14] Z. Tang, Y. Su, M.J. Er, F. Qi, L. Zhang, J. Zhou, A local binary pattern based texture descriptors for classification of tea leaves, Neurocomputing 168 (2015) 1011–1023.
[15] A. Xanthopoulou, I. Ganopoulos, A. Kalivas, M. Osathanunkul, P. Chatzopoulou, A. Tsaftaris, P. Madesis, Multiplex HRM analysis as a tool for rapid molecular authentication of nine herbal teas, Food Control 60 (2016) 113–116.
[16] P.H.G.D. Diniz, M.F. Barbosa, K.D.T. de Melo Milanez, M.F. Pistonesi, M.C.U. de Araújo, Using UV–vis spectroscopy for simultaneous geographical and varietal classification of tea infusions simulating a home-made tea cup, Food Chem. 192 (2016) 374–379.
[17] Q. Chen, J. Zhao, C.H. Fang, D. Wang, Feasibility study on identification of green, black and Oolong teas using near-infrared reflectance spectroscopy based on support vector machine (SVM), Spectrochim. Acta A 66 (2007) 568–574.
[18] K. Fraser, G.A. Lane, D.E. Otter, S.J. Harrison, S.Y. Quek, Y. Hemar, S. Rasmussen, Monitoring tea fermentation/manufacturing by direct analysis in real time