NEW STRATEGIES ON SAMPLE PREPARATION AND ON PLATFORMS DEVELOPMENT FOR TARGETED AND UNTARGETED METABOLOMICS ANALYSIS

MARÍA DEL MAR DELGADO POVEDANO
Supervisors: María Dolores Luque de Castro and Feliciano Priego Capote
Department of Analytical Chemistry. University of Córdoba
q82depom@uco.es

Metabolomics is defined as "the discipline that provides comprehensive and systematic information on temporal changes in the profiles of metabolites in biofluids and tissues, which can arise from control by the host genome, extended genomes and effect of other environmental or promoted factors". The vast field of metabolomics makes mandatory division into areas with limited content. A classification of metabolomics subdisciplines is based on the covered scope, which allows division into animal metabolomics (including clinical metabolomics) and plant metabolomics. Metabolomics analysis encompasses different strategies the nature of which depends on the objective of the study and previous knowledge of the biological problem [1]. The two most widely used strategies in metabolomics are untargeted and targeted analyses

The analytical process in metabolomics consists of four experimental parts: presampling and sampling, sample preparation, detection, and data treatment. Plant presampling depends on the type of system to be studied and sampling requires control of some aspects as cultivation conditions and part sampled. Animal presampling depends on cohort selection, and sampling depends on the kind of sample. Once the experimental steps have been established, the major current challenges in plant metabolomics analysis can be summarized as follows: (i) to establish presampling protocols, (ii) to improve the sample preparation and identification of unknown metabolites, (iii) to elucidate the metabolome of some plants; and (iv) to create new databases. While the major current challenges in animal metabolomics analysis are: (i) to use alternative samples, to improve the sampling and sample preparation protocols; (ii) to maximize the detection and identification capability; and (iii) to apply the developed methods to animal studies. Therefore, there are many challenges that need to be addressed in both plant and animal metabolomics.

In this sense, the current PhD Thesis entitled "New strategies on sample preparation and on platform development for targeted and untargeted metabolomics analysis" is focused on: 1) to contribute to plant metabolomics by opening new key research lines, and 2) to study in-depth an almost no used sample: sweat. Specific details on these topics are discussed below:

1. Vegetal metabolomics

This section includes: (i) the hydrolysis of oleuropein to its aglycon in olive leaf extracts accelerated by different hydrolases and ultrasound (US) action, (ii) the characterization of edible mushroom A. bisporus aqueous enzymatic extracts with antiviral activity against Hepatitis C virus (HCV), and (iii) the characterization of Cannabis sativa L. extracts.

1.1 Selective US-enhanced enzymatic hydrolysis of oleuropein to its aglycon in olive (*Olea europaea* L.) leaf extracts

Oleuropein and oleuropein aglycon, present in olive tree vegetal material and virgin olive oil, respectively, could be used as nutraceuticals or as components of new functional foods [2], due to their beneficial pharmacological properties [3,4]. Olive leaves are an ideal raw material to extract these compounds, and other valuable compounds they contain [5]. In this PhD Thesis, a fast US-assisted enzymatic hydrolysis method to obtain oleuropein aglycon from olive (*Olea europaea* L.) leaf extracts with minimum production of more degraded products was developed [6]. A drastic shortening of the time required for complete hydrolysis was achieved as compared with the traditional method based on enzymatic incubation, from 2 h to less than 20 min (**Figure 1).**

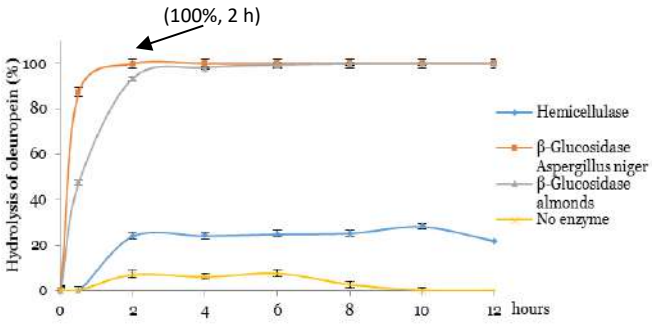

Figure 1. *Enzymatic effect of different enzymes on the hydrolysis of an oleuropein standard (10 µg/mL) as compared with the non-catalyzed reaction.*

The enzyme used in both cases was *β*-glucosidase from *Aspergillus niger (A. niger)* and the optimum US conditions for achieving maximum yield of oleuropein aglycon were 0.5 s/s duty cycle, 50 % amplitude and 45 s cycle, optimized by a Box-Behnken design, and 10 US cycles (in 18.75 min).

The method developed using oleuropein standard also allowed to obtain oleuropein aglycon from commercial and laboratory extracts from olive leaves. A previous extraction of oleuropein from olive leaf extracts was carried out by agitation with a methanol–water mixture, while hydrolysis was carried out by *β*-glucosidase from *A. niger*, and monitoring of degradation and quantitation was achieved by liquid chromatography–tandem mass spectrometry (LC–MS/MS) with multiple reaction monitoring. The hydrolysis yield of oleuropein was 90.8 % after 10 US cycles, which produced 90 % of oleuropein aglycon for commercial olive leaves extracts. The hydrolysis yield of oleuropein reached 95.7 % at 10 US cycles, with formation of 94.9% of the aglycon form for laboratory olive leaf extracts. The fast production of oleuropein aglycon thus achieved open a door to the exploitation of the interesting pharmacological properties of this compound [7], and its wide use for cosmetics and nutraceuticals [2].

1.2 Tentative identification of the composition of *Agaricus bisporus* aqueous enzymatic extracts with antiviral activity against HCV: A study by liquid chromatography—tandem mass spectrometry in high resolution mode

The "*in vitro*" inhibition of the protease NS3 of HCV by aqueous enzymatic extracts of the common edible mushroom *Agaricus bisporus* (*A. bisporus*) was recently described [8]. This fact suggests incorporation of these extracts to the diet as components of new functional foods or nutraceuticals for hepatitis-C prevention in people under risk of HCV infection. The use of these extracts as they are, as a part of new functional foods or as nutraceuticals, requires characterization of *A. bisporus* aqueous enzymatic extracts (AbAEE) as exhaustive as possible, with especial emphasis on checking their safety.

In this PhD Thesis also an LC–QTOF MS/MS platform was used to obtain, for the first time, the data for tentative identification of fifty five out of one hundred twenty-three total metabolites by combination of MS and MS/MS information on aqueous enzymatic extracts of the common edible mushroom *A. bisporus* [9], which had shown inhibition of the protease NS3 of HCV. The tentative identification of as many AbAEE components as possible was performed after an appropriate sample preparation strategy for fractionation based on sequential liquid–liquid extraction (LLE) of AbAEE with solvents of different polarity. The LC–QTOF MS/MS was used in high-resolution mode to obtain the required analytical information, which provided the necessary knowledge for a safe use of the extracts as they are very promising for HCV patients. Among the new identified compounds there are amino acids, sugars, carboxylic, fatty and cinnamic acids, mono- and disaccharides, phospholipids, and purines as the most outstanding; thus demonstrating that fractionation based on sequential LLE followed by LC–QTOF MS/MS is a suitable option to obtain a wide, representative snapshot of AbAEE composition.

1.3 Untargeted characterization of extracts from *Cannabis sativa* L. cultivars by chromatographic techniques coupled to mass spectrometry in high resolution mode

Cannabis is being explored increasingly for medicinal applications and therapies. It is used in the treatment of disorders such as migraine, spastic and pain disorders. Moreover, cannabinoids have been useful as antiemetic [10]. Countries such as Canada, the Netherlands and more than 50 % of the United States have authorized the medical application of *Cannabis* [11]. Most analytical studies on characterization of *Cannabis* plants were targeted at the detection of cannabinoids and terpenoids. The analysis of both families has been addressed with different analytical platforms [12], mainly by a combination of GC–MS and LC–MS [13,14]. There are scant studies dealing with compounds in *Cannabis sativa* other than cannabinoids and terpenoids [15]. Considering the interest on the study of *Cannabis sativa* in fields such as the pharmacological, biomedical, and agronomical more research on its composition is demanded to assess its potential for pharmacological use, but also to implement breeding programs with agronomical purposes.

In this PhD Thesis, a total of one hundred sixty-nine compounds (cannabinoids, terpenoids, lipids, flavonoids, among others, **Figure 2**) were identified in polar and no polar extracts from 17 cultivars of *Cannabis sativa* L. using the data obtained by GC–TOF/MS and LC–QTOF MS/MS platforms [16].

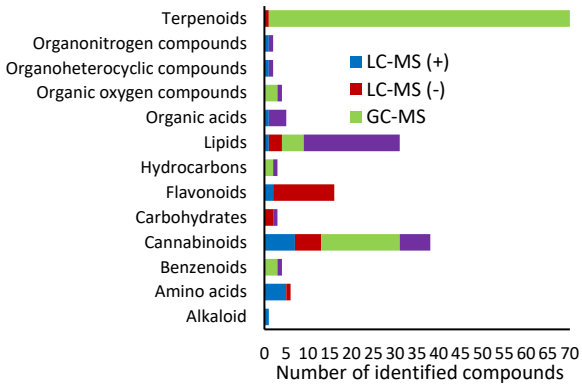

Figure 2. *Bars diagram comparing the number of compounds identified in Cannabis extracts by GC–TOF/MS, GC–TOF/MS after derivatization, LC–QTOF MS/MS in positive and negative ionization modes, classified by families.*

Relative contents of secondary metabolites (terpenoids and cannabinoids) in the same cultivars grown in greenhouse and in field were compared, and the found compositional differences between both types of grown conditions open a door for a more rational cultivation of this plant as a function of its use. The results of these studies revealed the suitability of both LC–QTOF MS/MS and GC–MS for characterization of polar/low-polar compounds from vegetal samples and provided the clues for identification of their metabolites.

2. Sweat and metabolomics: Sampling and sample preparation, and clinical applications

Sweat is mainly composed by water, but it contains several minor components including electrolytes, ammonia, urea, small molecules such as carboxylic acids and amino acids, and more complex biomolecules such as proteolytic enzymes and antimicrobial peptides, among others [17]. The varied composition, which can be modified by certain pathologies, and the no invasive sampling of this biofluid have gained clinical interest as a potential tool for diagnostics and biomarker monitoring. As a previous step to the use of this biofluid in the clinical field, metabolomics has been used as a tool to obtain a snapshot of the composition of sweat metabolome. A method developed for analysis of human sweat by LC–QTOF MS/MS allowed identifying forty-one compounds, mainly amino acids, the most abundant compounds in this sample [18]. From this research previously developed by the group of the PhD, two research lines emerged.

The first line consisted of using the platform previously proposed [19] to develop a more robust predictive model for lung cancer screening and a method for quantification of amino acids as potential biomarkers of lung cancer in the target biofluid by LC—MS/MS. The second line consisted of developing a GC—MS platform to elucidate sweat metabolome, then used to test different sample preparation protocols and samplings.

2.1 Clinical applications

Metabolomics had provided new technical approaches to enhance the capability of this discipline on biomarker discovery for diagnosis of different types of cancer. The potential of both sweat as biofluid and metabolomics for implementation in the diagnostic of lung cancer had been demonstrated in a previous study developed by the group of the PhD student [19]. Human sweat had been used as clinical sample to develop a tool for lung cancer screening. A prediction model based on a panel of sweat metabolites (80 % specificity and 79 % sensitivity), including amino acids, sugars, and some lipids, was built to discriminate patients with lung cancer from a control group with risk factor. These preliminary results, obtained with a unique cohort, emphasized the necessity of new studies to validate the obtained results and reduce the proportion of individuals that would be subjected to confirmatory tests.

In this PhD Thesis, a prediction model for lung cancer detection more robust than that obtained in the previous study was proposed [20]. The prediction model was based on two panels of sweat metabolites in which both false negatives and false positives were reduced in such a way that 95 % was the level reached for both specificity and sensitivity. The best panel providing specificity above 95 % (**Figure 3.A**) was formed by five metabolites: the monoglyceride MG(22:2), muconic, urocanic and suberic acids, and a tetrahexose.

This panel provided 96.9 % specificity and 83.8 % sensitivity. On the other hand, the best panel providing sensitivity above 95% (**Figure 3.B**) was composed of the metabolites included in the previous panel, in which suberic acid was replaced by nonanedioic acid (the previous panel provided 81.2 % specificity and 97.3 % sensitivity). The robustness of the previous biomarker panel was increased using new samples collected within different times (4 months and 2 years) and analyzed at different times (2012 and 2014, respectively) by different analysts, and always with the aim of discriminating lung cancer patients from smokers at risk factor.

A) **B)**

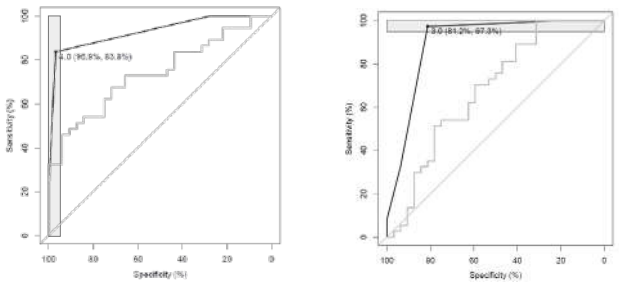

Figure 3. *ROC curves obtained from the panel with at least 95 % specificity (black) and that of nonanedioic acid (grey) as the best individual marker in terms of specificity (**A**) and ROC curves from the panel with at least 95 % sensitivity (black) and that of MG(22:2) (grey) as the best predictive marker in terms of sensitivity (**B**).*

Among 16 sweat metabolites used to discriminate lung cancer patients from control individuals, 5 of them were amino acids [19]. The composition of free amino acids in sweat is different from that in other biofluids, according to studies developed by automated analyzers [21], thus increasing the interest on this sample for quantitative analysis of amino acids for diagnosis purposes. Also, methods based on automated analyzers have detected up to 26 amino acids in human sweat, which allow confirming this biofluid as suitable for complete profile of amino acids [21]. However, derivatization of amino acids, a common practice prior to their photometric or fluorimetric detection, suffers from different drawbacks, which dramatically enlarge sample preparation protocols.

The development of a method to quantify 23 amino acids in human sweat by LC–MS/MS was afforded [22] in this section. Matrix effects decreased by both dilution of sweat and clean-up effect provided by centrifugal microsolid-phase extraction. Amino acids behave differently among them; therefore, compromise solutions were necessary for their determination in a single analysis. The concentration of amino acids in sweat ranges between 6.20 ng/mL (for homocysteine) and 259.77 µg/mL (for serine).

2.2 Sampling and sample preparation

The second line of the research on sweat-metabolomics was based on improving the detection coverage for identification of metabolites in this biofluid. Considering that more research on sweat composition is demanded to assess the potential of this biofluid for clinical diagnostic, and that before only NMR [23] and LC–MS [18] had been used for sweat characterization, sweat metabolome including volatile and easy to convert into volatile compounds needed to be elucidated.

This section of the PhD Thesis contains the development and validation of a method for analysis of human sweat by GC–TOF/MS in high resolution, which allowed studying a wide variety of these compounds [24]. GC–TOF/MS allowed detecting both volatile and no volatile compounds from human sweat after applying sweat deproteination and well-known derivatization reactions. It provided information for tentative identification of sixty-six compounds, including amino acids, dicarboxylic acids and other interesting metabolites such as myo-inositol or urocanic acid. GC–TOF/MS appears to be the best strategy to analyze this biofluid, in terms of metabolome coverage, as compared to previous studies using NMR and LC–MS (**Figure 4**).

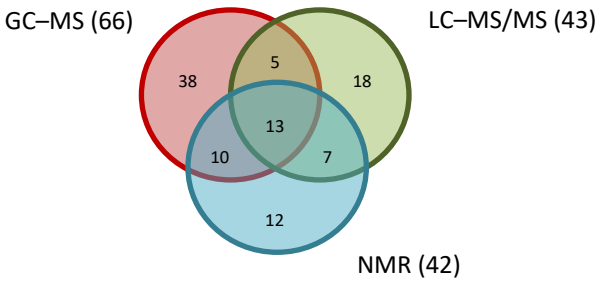

Figure 4. *Venn diagram comparing the number of compounds identified in sweat by NMR, LC–QTOF MS/MS and GC–TOF/MS.*

Among 66 compounds identified only 14 no polar compounds were detected, because most of them were of polar nature. Regarding no polar compounds, NMR analysis only allowed detection of some characteristic groups of lipids. The developed LC–QTOF MS/MS analysis confirmed that the main families of detected metabolites were polar compounds, and only 4 no polar compounds (3 fatty acids and 1 sphingolipid) were detected [18]. All the studies discussed above had in common the collection of the known as passive sweat — that collected from individuals at rest. The sampling devices employed could explain the differences in sweat composition, and the lack of no polar compounds [25]. Sweating is usually stimulated by heat and/or chemical induction (for instance, by pilocarpine), but also by exercise practice at different intensity levels.

Considering that sampling processes for this scantly known clinical sample in the above commented studies were based on passive sweat and that lipids had been scarcely detected in this sample, a method for active sweat analysis, with special emphasis on no polar compounds [26], was developed in this PhD Thesis. One hundred thirty-five compounds were tentatively identified by GC–TOF/MS. Lipids, volatile organic compounds, benzenoids and other interesting metabolites such as alkaloids and were identified — among them more than 40 lipids. Sample preparation to obtain a representative snapshot of active sweat metabolome consisted of methoximation plus silylation after LLE with dichloromethane. Compositional differences in passive and active sweat collected from the same volunteers were found, the latter presenting an enriched concentration of some key families of compounds such as fatty acids, alcohols, carbohydrates and no proteinogenic amino acids. The differences in concentration can be explained by the sampling protocol since sweat is stimulated in a different manner prior to collection of active or passive sweat.

Focusing again on sampling, all previous studies had in common the collection of the known as passive sweat but also that it was always collected as fresh sweat, and used as such [18,23,25]. In most of the wide variety of fresh sweat samplers it is mandatory to wear the device for sweat production, which is a tedious process. One alternative is collection of dry sweat. This protocol was proposed to standardize the sampling process due to its simplicity. A solid support as filter paper impregnated with 1:1 (v/v) ethanol–phosphate buffer to collect sweat components from the skin was proposed as a standardized protocol alternative to fresh sweat sampling [27]. The collected dry sweat allowed tentative identification of one hundred seventy-five compounds by using a combined approach based on GC–MS and LC–MS/MS analysis. The comparison between dry and fresh sweat revealed that the former provides an improvement in the detection of low-polar compounds, while fresh sweat is more suited for detection of polar metabolites. In addition, particular families such as carnitines, sphingolipids and N-acyl-amino acids never reported in fresh sweat were detected in dry sweat. Considering the variability sources affecting the sampling of fresh sweat, collection of dry sweat can be proposed as a standardized sample for its use in metabolomics and clinical studies.

Overall, this PhD Thesis has contributed to the Analytical Chemistry field in terms of developing methodologies, which help understanding the effect of sample preparation on detection coverage, and the favorable US effect on enzymatic reactions, as well as evaluating the effect of sampling and sample preparation on the detection of sweat components.

Acknowledgments:

I want to thank Profs. María Dolores Luque de Castro and Feliciano Priego Capote for being the supervisors of this PhD Thesis. I would also like to thank the MECD for my predoctoral FPU contract. I am greatly thankful to SEQA for receiving the prestigious Miguel Valcárcel award.

References:

[1] W.B. Dunn, D.I. Ellis, Metabolomics: current analytical platforms and methodologies, *Trends Anal. Chem.* 24 (2005) 285–294.
[2] I. Van der Stelt, E.F. Hoek-van den Hil, H.J.M. Swarts, J.J.M. Vervoort, L. Hoving, *et al.* Nutraceutical oleuropein supplementation prevents high fat diet-induced adiposity in mice, *J. Funct. Foods* 14 (2015) 702–715.
[3] I. Andreadou, E.K. Iliodromitis, E. Mikros, M. Constantinou, A. Agalias, *et al.* The olive constituent oleuropein exhibits antiischemic, antioxidative, and hypolipidemic effects in anesthetized rabbits, *J. Nutr.* 136 (2006) 2213–2219.
[4] A.R. Khalatbary, Gh. R. Zarrinjoei, Anti-inflammatory effect of oleuropein in experimental rat spinal cord trauma, *Iranian Red Crescent Med. J.* 14 (2012) 229–234.
[5] P.J. Xie, L.X. Huang, C.H. Zhang, Y.L. Zhang, Nutrient assessment of olive leaf residues processed by solid-state fermentation as an innovative feedstuff additive, *J. Appl. Microbiol.* 121 (2016) 28–40.
[6] M.D. Delgado-Povedano, F. Priego-Capote, M.D. Luque de Castro, Selective ultrasound-enhanced enzymatic hydrolysis of oleuropein to its aglycon in olive (*Olea europaea* L.) leaf extracts, *Food Chem.* 220 (2017) 282–288.
[7] F. Casamenti, C. Grossi, S. Rigacci, D. Pantano, I. Luccarini, *et al.* Oleuropein aglycone, a possible drug against degenerative conditions. In vivo evidence of its effectiveness against Alzheimer's disease, *J. Alzheimer's Dis.* 45 (2015) 679–688.
[8] P. Gayego-Yerga, J. Bautista, Trabajo fin de grado, Sevilla (Spain): Facultad de Biología, Universidad de Sevilla, 2015.
[9] M.M. Delgado-Povedano, V. Sánchez de Medina, J. Bautista, F. Priego-Capote, M.D. Luque de Castro, Tentative identification of the composition of Agaricus bisporus aqueous enzymatic extracts with antiviral activity against HCV: A study by liquid chromatography–tandem mass spectrometry in high resolution mode, *J. Funct. Foods* 24 (2016) 403–419.
[10] F. Grotenhermen, Cannabis and cannabinoids, pharmacology, toxicology, and therapeutic potential, Ethan Russo, 2002.
[11] R. Abuhasira, L. Shbiro, Y. Landschaft, Medical use of Cannabis and cannabinoids containing products–Regulations in Europe and North America, *Eur. J. Intern. Med.* 49 (2018) 2–6.
[12] A. Leghissa, Z.L. Hildenbrand, K.A. Schug, A review of methods for the chemical characterization of cannabis natural products, *J. Sep. Sci.* 41 (2018) 398–415.
[13] A. Leghissa, Method development for qualification and quantification of cannabinoids and terpenes in extracts by gas chromatography–mass spectrometry. The University of Texas at Arlington, Arlington, 2016.
[14] A. Leghissa, Z.L. Hildenbrand, F.W. Foss Jr., K.A. Schug, Determination of cannabinoids from a surrogate hops matrix using multiple reaction monitoring gas chromatography–triple quadrupole-mass spectrometry, J. Sep. Sci. 2018, https://doi.org/10.1002/jssc.201700946
[15] F. Ahmad, T. Abbas, K. Farman, A. Akrem, M.A. Saleem, *et al.* High-throughput phytochemical characterization of non-cannabinoid compounds of Cannabis plant and seed, from Pakistan, *Pak. J. Bot.* 50 (2018) 639–643.
[16] M.M. Delgado-Povedano, C. Sánchez-Carnerero Callado, F. Priego-Capote, C. Ferreiro-Vera, Untargeted characterization of extracts from *Cannabis sativa* L. cultivars by gas and liquid chromatography coupled to mass spectrometry in high resolution mode, *Talanta* 208 (2020) 120384.
[17] B. Schittek, R. Hipfel, B. Sauer, J. Bauer, H. Kalbacher, *et al.* Dermcidin: a novel human antibiotic peptide secreted by sweat glands, *Nat. Immunol.* 2 (2001) 1133–1137.
[18] M. Calderón-Santiago, F. Priego-Capote, B. Jurado-Gámez, M.D. Luque de Castro, Optimization study for metabolomics analysis of human sweat by liquid chromatography–tandem mass spectrometry in high resolution mode, *J. Chromatogr. A* 1333 (2014) 70–78.
[19] M. Calderón-Santiago, F. Priego-Capote, N. Turck, X. Robin, B. Jurado-Gámez, *et al.* Human sweat metabolomics for lung cancer screening, *Anal. Bioanal. Chem.* 407 (2015) 5381–5392.
[20] M.M. Delgado-Povedano, M. Calderón-Santiago, F. Priego-Capote, et al. Recent advances in human sweat metabolomics for lung cancer screening, *Metabolomics* 12 (2016). https://doi.org/10.1007/s11306-016-1116-4
[21] H. Mark, C.R. Harding, Amino acid composition, including key derivatives of eccrine sweat: potential biomarkers of certain atopic skin conditions, *Int. J. Cosmetic Sci.* 35 (2013) 163–168.
[22] M.M. Delgado-Povedano, M. Calderón-Santiago, F. Priego-Capote, M.D. Luque de Castro, Study of sample preparation for quantitative analysis of amino acids in human sweat by liquid chromatography-tandem mass spectrometry, Talanta 146 (2016) 310–317.
[23] M. Harker, H. Coulson, I. Fairweather, D. Taylor, C.A. Daykin, Study of metabolite composition of eccrine sweat from healthy male and female human subjects by ^{1}H NMR spectroscopy, *Metabolomics* 2 (2006) 105–112.
[24] M.M Delgado-Povedano, M. Calderón-Santiago, F. Priego-Capote, M.D. Luque de Castro, Development of a method for enhancing metabolomics coverage of human sweat by gas chromatography-mass spectrometry in high resolution mode, *Anal. Chim. Acta* 905 (2016) 115–125.
[25] V.P. Kutyshenko, M. Molchanov, P. Beskaravayny, V.N. Uversky, M.A. Timchenko, Analyzing and mapping sweat metabolomics by high resolution NMR spectroscopy, PLoS ONE 6 (2011) e28824.
[26] M.M. Delgado-Povedano, M. Calderón-Santiago, M.D. Luque de Castro, F. Priego-Capote, Metabolomics analysis of human sweat collected after moderate exercise, *Talanta* 177 (2018) 47–65.
[27] M.M. Delgado-Povedano, L.S. Castillo-Peinado, M. Calderón-Santiago, M.D. Luque de Castro, F. Priego-Capote, Dry sweat as sample for metabolomics analysis, *Talanta* 208 (2020) 120428.