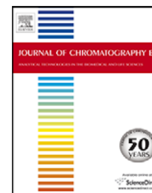
Journal of Chromatography B, 871 (2008) 253–260

Contents lists available at ScienceDirect

Journal of Chromatography B

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



Global metabolic profiling analysis on human urine by UPLC–TOFMS: Issues and method validation in nutritional metabolomics☆

Philippe A. Guy*, Isabelle Tavazzi, Stephen J. Bruce, Ziad Ramadan, Sunil Kochhar

BioAnalytical Science Department, Nestlé Research Center, P.O. Box 44, Vers-chez-les Blanc, 1000 Lausanne 26, Switzerland

ARTICLE INFO

Article history:
Received 1 February 2008
Accepted 13 April 2008
Available online 29 April 2008

Keywords:
Metabolomics
Metabonomics
UPLC–TOFMS
Urine
Method validation

ABSTRACT

Optimisation and method validation was assessed here for metabolic profiling analysis of urine samples using UPLC–TOFMS. A longer run time of 31 min revealed greater reproducibility, and the higher number of variables was identified as compared to shortened run times (10 and 26 min). We have also implemented two QC urine samples enabling the assessment of the quality and reproducibility of the data generated during the whole analytical workflow (retention time drift, mass precision and fluctuation of the ion responses over time). Based on the QC data, suitable standards for ensuring consistent analytical results for metabolomics applications using the UPLC–MS techniques are recommended.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Metabolomics research is commonly performed in the context of systems biology [1,2], which involves integrated study of the transcriptome, proteome, and metabolome. Metabolomics is viewed as a potentially powerful application in many fields, for example, in specific biomarker discovery for clinical diagnostics, drug discovery, or nutritional intervention based on personalised nutrition applications.

The metabolome can be defined as the entire cellular complement of endogenous low molecular weight (typically <1000 Da) biomolecules [3]. It includes, but is not limited to, amino acids, carbohydrates, lipids, peptides, purines, pyrimidines, vitamins, and numerous metabolites involved in biosynthesis, biodegradation pathways and serving various functions [e.g., antioxidants, cofactors, intra- and inter-cellular regulatory and signalling molecules (i.e., hormones, enzyme inhibitors, neurotransmitters, etc.)]. Bioanalytical techniques used for metabolomics should thus be capable of accurately monitoring numerous known and unknown molecules that span a diverse chemical spectrum and large dynamic concentration range (estimated to be 7–9 orders of magnitude, ranging from pmol to mmol) [4]. Several analytical techniques are commonly used for metabolomics applications including mainly ^{1}H NMR and/or MS-based approaches. NMR has demonstrated its great potential essentially due to a high reproducibility of measurements and throughput of analysis [5,6]. Moreover, appropriate databases have been progressively established [7–13]. However, a major problem of NMR metabolomic studies relates to the relatively poor sensitivity of the technique (μg range). In parallel, GC, CE or LC separation techniques coupled to on-line MS detection have been widely accepted as complementary techniques to NMR, essentially in terms of sensitivity. Recognising that no single technique can be expected to meet all of the field's diverse challenges, many metabolomics research programs employ several analytical techniques.

Metabolomic studies are very difficult to perform because of the wide variability and flexibility of the biofluids (mainly urine) associated with different confounding factors such as gender, age, time of day, health state, lifestyle, diet, phenotypes [14]. In contrast to NMR spectroscopy, reproducibility of LC or ultra high pressure liquid chromatography (UPLC) coupled to MS is more challenging and needs to be precisely assessed. Indeed, the ability to demonstrate that high quality data is obtained is an absolute prerequisite to support any sound biological insights. Although metabolomic studies require a high analytical quality standard, this is even more important for nutritional applications where subtle changes in metabolic profiles are expected as compared to toxicological applications. In 2005, an initiative was undertaken to

☆ This paper is part of a special volume entitled "Hyphenated Techniques for Global Metabolite Profiling", guest edited by Georgios Theodoridis and Ian D. Wilson.

* Corresponding author. Tel.: +41 21 785 94 37; fax: +41 21 785 94 86.
E-mail address: philippe.guy@rdls.nestle.com (P.A. Guy).

1570-0232/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.jchromb.2008.04.034

propose some recommendations for standardisation and reporting of metabolic analyses [15]. Surprisingly, only a few publications report factors affecting the method performance for urine samples using LC–MS or UPLC–MS techniques. Waybright et al. have focused their work on the optimisation of urine sample preparation and analytical parameters, but the reproducibility aspect of their data was focused on retention times only [16]. Wagner et al. described a targeted metabolites approach for measuring mercapturic acids in human urine assessed by replicate analyses of a randomly chosen sample [17,18]. The within-day LC–MS data reproducibility for metabolomics analysis was reported by Gika et al. from human urine samples. The authors emphasised the use of QC data to validate datasets before any further statistical treatment and have introduced a workflow acceptance criteria [19]. Finally, we found only one publication describing a quality control strategy for metabolomics analysis using UPLC–MS [20]. The present work describes the influence of the chromatographic conditions on the reproducibility and the number of ions detected by UPLC–TOFMS. Moreover, we have implemented two QC urine samples to assess the good quality data for metabolomic study and to monitor the performance of the instrument over time. The focus of the present study was on the method optimisation and validation side rather than to identify putative biomarkers.

2. Experimental

2.1. Reagents

Acetonitrile containing 0.1% formic acid (v/v) and water containing 0.1% formic acid (v/v), and water ULC/MS grade were purchased from Biosolve (Brunschwig, Basel, CH). Leucine enkephalin was purchased from Sigma (Buchs, CH). Methanol HPLC grade was purchased from Fisher Scientific (Wohlen, CH). Metabolomics performance test mix (Waters, Milford, MA, USA) comprised a mixture of hippuric acid, theophylline, caffeine, and nortriptyline. Isotopically labelled internal standard d_5-phenylalanine (ring-d_5, isotopic purity 98%) was purchased from Cambridge Isotope Laboratories (Baumgarten, Switzerland).

2.2. Sample preparation

Urine samples were collected from different volunteers and aliquoted into 1.5 mL Eppendorf tubes before storage at −80 °C. Urine aliquot samples were thawed at room temperature and 25 μL was introduced into an amber LC–MS certified vial containing another 75 μL distilled water ULC/MS grade (1/4 dilution). The vial was flushed under a nitrogen stream, sealed and placed in the autosampler before analysis. A 5 μL volume of the diluted urine was injected and analysed by UPLC–TOFMS.

2.3. UPLC conditions

Separation of the metabolites was carried out using the Waters Acquity UPLC system equipped with the Acquity BEH C18 column (2.1 mm × 100 mm, 1.7 μm). Mobile phases were composed of (A) water containing 0.1% formic acid and (B) acetonitrile/water (80:20, v/v) containing 0.1% formic acid. The analytical column and the autosampler were maintained at a temperature of 30 °C (±2 °C) and 10 °C (±4 °C), respectively. A constant flow rate of 400 μL/min was used during the whole analysis, giving a typical back pressure of 7200–11000 psi (ca. 500–760 bars). The gradient used was the following: from 0 to 3 min, 1% B; 3 to 18 min, 22% B; 18 to 30 min, 50% B; 30 to 31 min, 99% B; 31 to 36 min, 99%B, 36 to 36.1 min, 1% B; 36.1 to 40 min, 1% B. The eluant was directed from the UPLC to the mass spectrometer between 0.5 and 31 min, and to waste at the beginning and end of the gradient.

2.4. MS conditions

The MS detection was obtained on an LCT Premier TOF mass spectrometer (Waters, Milford, MA) equipped with an electrospray ionisation source and lockspray interface for accurate mass measurements. MS source parameters were set with capillary, aperture 1 and cone voltages of 2500, 5 and 40 V, respectively. The desolvation gas and source temperatures were set at 250 °C and 120 °C, respectively, whereas the nebuliser and cone gases flow were respectively set at 700 L/h and 3 L/h. Data acquisition was realised (with dynamic range enhancement mode activated) both in negative and positive centroid acquisition mode (W analyser mode) using MassLynx software version 4.1 (Waters, Milford, MA). Full scan mass acquisition was performed by scanning an m/z range of 100–1000 Th in a 0.2 s per scan (with an inter-scan time delay of 0.01 s). Lock mass calibration was realised by infusing a solution of Leu-enkephalin (1 μg/mL solubilised in acetonitrile/water (1:1, v/v) containing 0.1% formic acid) into the MS instrument every 20 scans (5 scans were averaged). The protonated molecule (m/z 557.2802) and its attenuated ion (m/z 556.2771) were used for positive ionisation mode (m/z 555.2645 and 554.6233 for negative ionisation mode).

2.5. QA/QC samples

A large quantity of in-house urine, distinct from the samples to be measured for a clinical study, was used as the QC1 sample. Numerous aliquots were then prepared and kept at −80 °C until use. The metabolomic Waters performance kit standards were solubilised with 1000 μL water/acetonitrile (98:2, v/v) to provide the corresponding concentrations of 120 ng/μL (theophylline, caffeine and hippuric acid), 60 ng/μL (4-nitrobenzoic acid), and 45 ng/μL (nortriptyline). A 20 μL volume of this mixture was aliquoted into different LC–MS vials, flushed under nitrogen and stored at −30 °C until use. Every day, an aliquot was thawed at room temperature, and to this 55 μL water followed by 25 μL pooled urine were added, flushed under nitrogen and placed in the autosampler.

In addition, an aliquot of 50 μL of each urine sample belonging to a clinical study was pooled to provide the QC2 sample. From this pooled urine, several aliquots were prepared and kept at −80 °C until use. Every day, a fresh aliquot QC2 was thawed and processed as a normal sample.

2.6. Batch analysis

When starting a new metabolomics project, a dedicated new UPLC column is used and equilibrated after running roughly 25 blanks under the chromatographic conditions described above. The typical batch sequence of urine samples consisted in the consecutive analysis of 2 QC1 spiked urine samples, 1 QC2 pooled urine sample (at the beginning of the study or the week), followed by 1 QC1 spiked urine sample, 1 QC2 pooled urine sample, 10 unknown urine samples, 1 QC1 spiked urine sample, 1 QC2 pooled urine sample, before running 10 unknown urine samples, etc. An identical sequence was repeated to complete the total set of samples ($n = 77$, including QCs) analysed in less than 3 days per ionisation mode. At the end of the last QC2 pooled urine sample, the chromatographic system was flushed with 10% acetonitrile in water for 2 h running at a flow rate of 0.2 mL/min. Moreover, the MS interface was cleaned (with methanol and water) once a week to ensure consistent and reliable results and to prevent any tedious instrument maintenance in the long term.

2.7. Data treatment

QC1 urine samples: retention time (RT) consistency, accurate mass precision and ion intensities (peak area and height) were confirmed throughout the runs by assessing the spiked analytes (theophylline, caffeine, nortriptyline, hippuric acid for positive ionisation mode, and hippuric acid, citrate, succinate for negative ionisation mode). This information was automatically extracted from the raw data using QuanLynx software (Waters, Milford, USA). QC2 pooled urine samples: These runs along with the unknown urine samples were pre-processed using MarkerLynx™ application manager software version 4.1 (Waters, Milford, USA) to extract and align RT and mass precision variables (calculated as mean values) and to normalise the ion intensities. This dataset was then exported as a txt file and further processed using SIMCA-P+ software version 11 (Umetrics, Umea, Sweden).

3. Results and discussion

As a global "holistic" metabolic fingerprinting profile, sample handling before LC–MS analysis needs to be simplified as much as possible to avoid any bias results. One-to-four dilution step (in water) of urine samples prior to UPLC–MS measurement does fulfil this criterion. To ensure the consistency of our results, we have analysed different urine samples, originating from a human lifestyle study ($n = 10$ samples), using various chromatographic conditions.

3.1. Optimisation of the UPLC chromatographic conditions

Several authors have reported the use of fast gradient conditions (12-min total run time) [21–24] and even faster ones (1.5-min total run time) [25] for urine sample metabolic profiling, thanks to the UPLC separation power. In this work, we have assessed the quality of data generated from various gradient conditions using the UPLC system coupled to the LCT Premier MS detector. Two different sets of urine samples originating from males and females ($n = 5$, each) were analysed. Each urine sample was injected in duplicate using three UPLC conditions (i.e., 0–10 min for gradient 1, 0–26 min for gradient 2, and 0–31 min for gradient 3). Fig. 1 depicts the typical total ion chromatogram obtained for the 10-, 26- and 31-min gradients. Under these conditions, typical peak widths of 6, 17, and 24 s were obtained, respectively. The shortened 10-min linear gradient (Fig. 1a) depicts a poor separation of analytes from 0.5–2 to 4–7.5 min, as compared to the other gradients. On the other hand, the 31-min (three steps linear) gradient revealed a better chromatographic separation compared to 10-min and 26-min gradient conditions, especially in the second half of the run (Fig. 1c).

3.2. Data handling

Peak picking, alignment (RT and *m*/*z* values) and ion intensity normalisation were realised using MarkerLynx™ application manager software. The peak detection parameters set in MarkerLynx™ were peak to peak baseline noise of 200, peak width of 12 s (at 5% height), noise elimination level of 30, deisotope filtering activated, running from 0 to 42 min. From each gradient condition, RT, masses and significance information obtained underwent for further statistical data treatment using SIMCA-P+ software. Fortunately, similar variable pairs (RT and mass) were observed for the three gradient conditions, but additional ones were also highlighted with the longer run time (i.e., 31-min gradient). Indeed, the total RT/mass pairs obtained for each gradient condition were 2378 for the 10-min gradient, 8743 for the 26-min gradient, and 21,862 for the 31-min gradient. Obviously, the data handling was performed under identical conditions, reflecting that the difference observed is related to the gradient conditions. Consequently, the number of variables was increased by a factor of 3.6-fold (26-min gradient), and 9.2-fold (31-min gradient) as compared to the 10-min gradient. The additional ions observed in longer run time (gradient 2 and 3) are evenly distributed through the chromatogram. Obviously, these new variables can reflect both additional metabolites, chemical noise and possibly new adduct ions. However, the latter possibility is less obvious as same HPLC solvents and MS parameters were used for the three experiments. However, due to a better chromatographic resolution of the compounds for longer run time, some of these adducts may be visible. These results may also indicate possible ion suppression related to shortened run time, due to co-eluting molecules. Indeed, the ionisation efficiency of low abundant molecules will be affected (mainly suppressed) by the high abundant ones. Nordström et al. have also shown that the length of separation was important for the number of detected features [26]. Indeed, they have analysed identical human serum spiked or not with a mixture of endogenous compounds. Their UPLC–TOFMS analyses realised over a 31-min run time provided 30% more variables compared to a 10-min run time. However, the impact of UPLC is indisputable due to overall improved performance, where the peak capacity and the number of marker ions detected using a fast UPLC separation (1.5 min) was found to be similar to that generated by conventional HPLC methods with a 10-min separation [25].

Therefore, based on these results, two analytical strategies are possible: (1) reducing the analytical run time to enable a higher throughput of analysis or (2) increasing the analytical run time to increase the number of molecules detected with an overall lower throughput. Another solution to get both high throughput and good data quality would be to use the latest MS instrumentation, combining high pressure chromatography with ion mobility separation on the MS side (Synapt HDMS Q-TOF instrument). Indeed, an additional separation dimension of molecules due to their characteristic size and shape could improve the co-elution issue, while maintaining high throughput analysis. Another interesting approach would be to use 2D LC separation up-front before MS detection using high pressure chromatography.

3.3. Repeatability (intra- and inter-assays)

From the dataset generated, principal component analysis (PCA, an unsupervised method) was realised using mean centred and Pareto scaling. This latter scaling approach was used to avoid modelling chemical noise, as the square root of the standard deviation is used. Fig. 2 depicts the PCA obtained for the 10 urine samples analysed in duplicate under the three chromatographic conditions tested. Among the three PCA plots, the most reproducible results were obtained under the 31-min gradient conditions (Fig. 2c). Indeed, the two replicates are very closely related in the PCA plot, as compared to the other two gradients evaluated. Although it is obvious to see the difference in repeatability of each gradient by eye, we have calculated the median of the two-dimensional Euclidean distance (ED) for each PCA model. For two 2D points ($P = (p_{\text{pc1}}, p_{\text{pc2}})$ and $Q = (q_{\text{pc1}}, q_{\text{pc2}})$), the distance is computed as:

$$\sqrt{(p_{\text{pc1}} - q_{\text{pc1}})^2 + (p_{\text{pc2}} - q_{\text{pc2}})^2},$$

where pc1 and pc2 are the first and second principal component scores. Therefore, the ED values calculated for each gradient tested were ED = 35.2 (gradient 1), ED = 20.7 (gradient 2), and ED = 3.9 (gradient 3). A lower ED value (gradient 3) reflects a better repeatability, as compared to the other two gradients. In addition, we calculated the variance captured by each principal component (PC) to provide

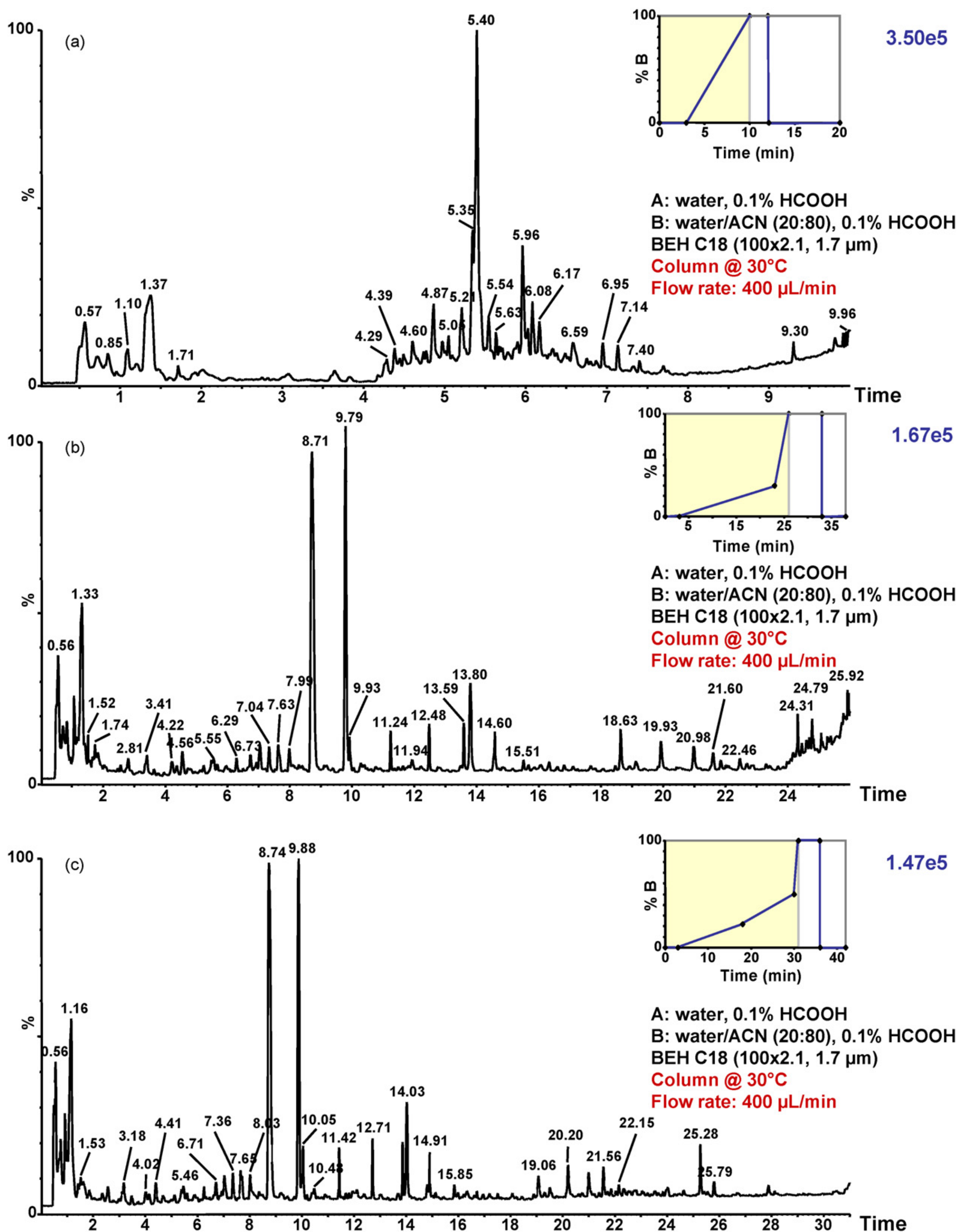


Fig. 1. Typical chromatograms obtained from the analysis of a male urine sample analysed by UPLC–TOFMS in positive ESI mode. The UPLC gradient was realised within (a) a 10-min, (b) a 26-min, and (c) a 31-min gradient conditions. Obviously, all other analytical parameters were kept the same. The gradient conditions used for the three experiments are reported as insert.

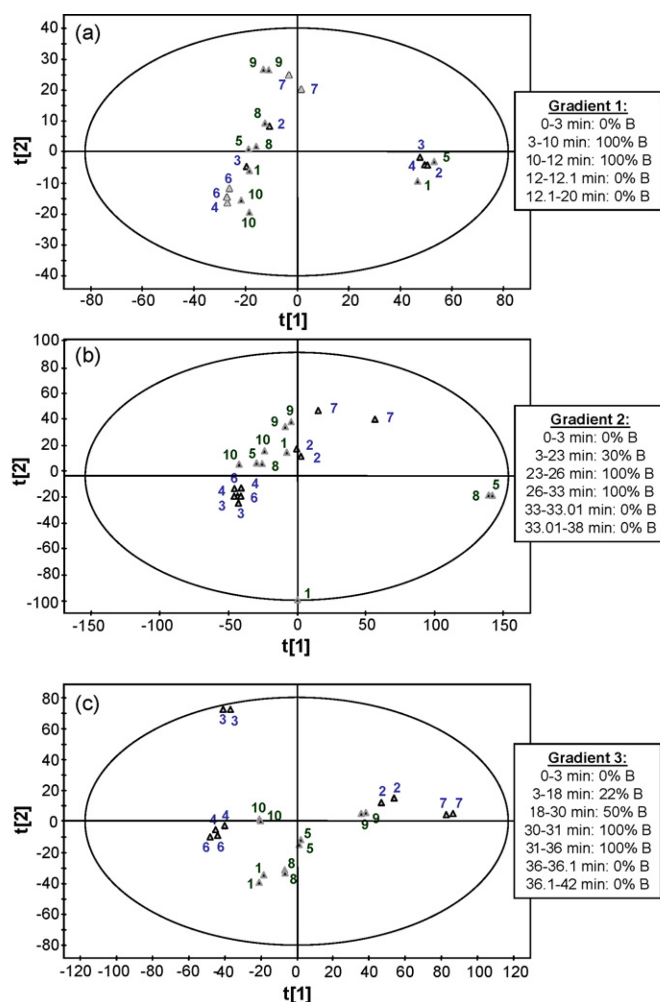


Fig. 2. Principal component analysis (PCA, mean centred and Pareto scaling) of duplicate analysis of urine samples ($n = 5$ for males and females) obtained from the (a) 10-min, (b) 26-min, and (c) 31-min gradient conditions. Male and female urine samples are depicted in grey and black triangles, respectively.

figures on the variation due to the repeatability of each chromatographic method. The percent of variations captured by each gradient were pc1 = 42%, pc2 = 11% (gradient 1), pc1 = 33%, pc2 = 11% (gradient 2), and pc1 = 24%, pc2 = 11% (gradient 3). From the score plot, we can see that pc1 is related to the variations between samples (higher in gradient 1 due to low repeatability). The loading plot was not locked because we were not interested to know what was causing these variations other than knowing the number of metabolites measured. The low repeatability may be explained by the number of data-points acquired for the three gradients. Thus, careful observations on the results generated by the three gradients revealed roughly 5 (10-min gradient), 13 (26-min gradient), and 27 (31-min gradient) data-points to define a peak. Considering that a minimum of 10 data-points is normally required to provide a well-defined peak to achieve reproducible results, the shortened gradients evaluated may have hampered the quality of the data. Taking into account all the above results, the 31-min gradient conditions were kept for the analysis of urine samples. Additional experiments were then realised by injecting the same urine sample (i) seven times on the same day (intra-assay) and (ii) over a 7-day period (inter-assay ($n = 4$), data not shown). Good repeatability (ion response) for both intra- and inter-assay measurements was obtained using the 31-min gradient with respective CV values calculated at 4% and 12% (for theophylline), at 9% and 9% (for caffeine), at 10% and 21% (for hippuric acid), and at 12% and 8% (for nortriptyline).

3.4. System suitability test

Implementation of quality control checks was evaluated over a relatively small lifestyle project in which 20 healthy volunteers were submitted to an identical diet. Sampling of urine was realised on three separate occasions ($n = 3$ for each volunteer), giving a total of 60 urine samples. UPLC–TOFMS analyses were carried out both in positive and negative ionisation acquisition modes, using QC1 and QC2 urine samples (as described in Section 2). According to the 31-min gradient, the total set of urine sample measurements were completed within 6 days in electrospray positive and negative ionisation modes ($n = 77$ runs for each acquisition mode). After acquiring the whole set of urine samples, RT along with peak area and height observed for the seven QC1 samples were extracted using QuanLynx software. Table 1 summarises the data obtained for each of the spiked compounds both in positive and negative ionisation modes. RT and mass precision shifts were below 0.03 min and 5.0 ppm (mean value) for positive ionisation mode (and below 0.03 min and 4.6 ppm for negative ionisation mode). In addition, elemental composition realised from a small number of scans at the peak apex resulted in the expected composition hit in terms of good mass accuracy and isotopic pattern (i-fit value ranking in the first position in all cases). The latter value is a measure of likelihood that the isotopic pattern of the elemental composition matches a cluster of peaks in the spectrum. Thus, a lower value of i-fit, in addition to a good mass accuracy, reveals a better chance to assign the correct elemental formula proposed by the software. Among the various parameters checked, fluctuation of the ion response was the most critical one, as already reported by several authors [19,20,27]. Coefficients of variation (CV) of the area measured for the four spiked compounds ranged from 3% (hippuric acid) up to 22% (nortriptyline) in positive ionisation mode, and from 7% (hippuric acid) up to 16% (succinate) in negative ionisation mode. In addition, the ratios of peak height against peak area provide an indication of the good chromatographic performance (i.e., identical peak width). Indeed, an overall decrease in these ratios will reveal a deterioration of the UPLC column due to peak broadening, which was not the case for our measurements. Obviously, a clear cut-off decision is rather difficult to make. A QC1 sample judged to be outside the acceptance criteria will result in exclusion of all unknown samples measured before and after the “bad” QC1, and will need to be re-analysed. In the general guidance defined by the FDA documentation, the precision determined at each of the concentration levels should not exceed a CV calculated above ±15% (and ±20% when reaching the limit of quantification) [28]. This guidance also mentions that 2 out of 6 QC samples (i.e., 33%) can fall outside the former acceptance criteria, while the whole sequence will still be accepted. This document was not established for metabolomics applications but more for targeted approaches (with available standard compounds). In such cases, an optimised sample preparation must be developed, leading to less ion suppression/enhancement phenomena due to “clean” extracts. On the contrary, urine samples may differ greatly between individuals and also due to the time-point of sampling. With our belief and experience when dealing with metabolomic analysis of urine samples, we propose that a CV value greater than ±25% should be considered as outside the acceptance window. However, if one QC1 sample falls outside this value, then all the unknown urine samples analysed before and after the QC1 excluded (until the next valid one) will need to be re-analysed.

All these quality control checks are vital before pre-processing the samples of any study. Indeed, some key parameters need to

Table 1
Quality assessment of QC1 runs ($n = 7$, analysed every 10 urine samples) obtained from UPLC–TOFMS positive and negative ionisation modes

QC1 (run number)	1	2	3	4	5	6	7	Mean	CV
ESI (+)									
Theophylline $C_7H_8N_4O_2$									
RT (min)	7.80	7.81	7.81	7.81	7.81	7.80	7.83	7.81	0.13%
Area	275	273	291	289	295	291	265	283	4%
Height	3,330	3,290	3,290	3,250	3,290	3,320	3,310	3,297	1%
Height/area	12.1	12.1	11.3	11.3	11.1	11.4	12.5	11.7	5%
Mass error (ppm)	−4.4	−5.0	−7.7	−6.7	−3.9	3.9	−2.2	−3.7	
i-Fit	0	1	2	1	4	2	2	2	
Hippuric ac. $C_9H_9NO_3$									
RT (min)	8.84	8.85	8.86	8.85	8.85	8.85	8.86	8.85	0.08%
Area	338	355	336	355	334	349	324	342	3%
Height	3,060	3,300	3,240	3,240	3,080	3,120	3,220	3,180	3%
Height/area	9.1	9.3	9.6	9.1	9.2	9.0	9.9	9.3	4%
Mass error (ppm)	−12.8	−7.2	−6.1	5.6	−9.4	10.6	−2.2	−3.1	
i-Fit	3	1	1	30	3	35	1	10	
Caffeine $C_8H_{10}N_4O_2$									
RT (min)	10.18	10.17	10.18	10.18	10.18	10.18	10.19	10.18	0.06%
Area	312	357	344	361	345	362	295	340	8%
Height	4,290	5,040	4,610	4,820	4,610	4,750	3,740	4,551	9%
Height/area	13.8	14.1	13.4	13.3	13.4	13.1	12.7	13.4	3%
Mass error (ppm)	5.6	2.6	4.1	−4.6	−4.1	0.5	−2.1	0.3	
i-Fit	83	72	94	70	6	81	50	65	
Nortriptyline $C_{19}H_{21}N$									
RT (min)	25.85	25.84	25.83	25.84	25.83	25.83	25.82	25.83	0.04%
Area	290	398	477	456	488	401	603	445	22%
Height	3,320	4,610	5,570	5,800	6,090	4,750	7,080	5,317	23%
Height/area	11.4	11.6	11.7	12.7	12.5	11.8	11.7	11.9	4%
Mass error (ppm)	−7.6	−7.9	4.5	−1.9	−12.1	−6.8	−3.4	−5.0	
i-Fit	2	237	398	394	6	333	367	248	
QC1 (run number)	**1**	**2**	**3**	**4**	**5**	**8**	**9**	**Mean**	**CV**
ESI (−)									
Citrate $C_6H_8O_7$									
RT (min)	1.10	1.10	1.10	1.11	1.10	1.08	1.08	1.10	1.03%
Area	4,004	3,779	3,617	3,664	3,288	3,106	2,787	3,463	12%
Height	31,241	29,191	28,923	29,342	24,077	22,332	20,304	26,487	16%
Height/Area	7.8	7.7	8.0	8.0	7.3	7.2	7.3	7.6	5%
Mass error (ppm)	−9.4	3.1	5.8	−5.2	−13.1	−4.7	−6.3	−4.3	
i-Fit	11.2	16.3	18.3	25.3	35.8	5.3	16.3	18.4	
Succinate $C_4H_6O_4$									
RT (min)	1.51	1.51	1.51	1.51	1.51	1.51	1.50	1.51	0.25%
Area	11.5	8.2	10.0	11.7	9.3	9.6	7.5	9.7	16%
Height	130	109	133	139	115	115	98	120	12%
Height/Area	11.3	13.3	13.3	11.9	12.3	12.0	13.1	12.5	6%
Mass error (ppm)	1.7	1.7	−0.9	−3.4	2.6	−6.8	0.6	−0.6	
i-Fit	3.3	4.7	2.2	2.8	3.6	4.4	1.5	3.2	
Hippuric ac $C_9H_9NO_3$									
RT (min)	8.79	8.79	8.80	8.80	8.79	8.78	8.79	8.79	0.08%
Area	5,013	4,920	4,939	4,873	4,574	4,445	4,139	4,700	7%
Height	43,123	42,341	43,160	42,690	39,600	38,570	35,887	40,767	7%
Height/Area	8.6	8.6	8.7	8.8	8.7	8.7	8.7	8.7	1%
Mass error (ppm)	2.8	0.0	0.0	6.2	−4.5	0.6	−0.6	0.6	
i-Fit	43.9	15.3	15.5	42.3	6.5	21.3	20.8	23.7	

be defined, within MarkerLynx™ software, in order to adequately align both RT and observed mass. Fig. 3 depicts the PCA score plots obtained from all unknown urine samples plus the QC2 ones, after MarkerLynx™ pre-processing treatment (alignment and ion intensity normalisation). As one would guess, QC2 plots need to be tightly grouped together to ensure the consistency of the data (both within- and between-days). Although this was the case for the runs performed in negative ESI mode, a different behaviour was observed for QC2_6 and QC2_7 in positive ESI mode (Fig. 3a bottom). Therefore, the unknown urine samples analysed in between these two QCs along with two new QC runs (QC2_8 and QC2_9) were re-analysed. The loadings of these new QC2 runs fitted well with the former QC2 ones (for both PC axes). These examples highlight the importance of monitoring QC2, as they can be processed as the unknown samples, to ensure both good data quality obtained from the analytical side and to confirm that no bias is introduced during the statistical treatment.

3.5. Linearity and quantification aspects

To assess the suitability for quantitative or semi-quantitative measurements using UPLC–TOFMS, a calibration curve was constructed from a urine sample spiked with a mixture of caffeine, nortriptyline, hippuric acid, leucine, theophylline and d_5-phenylalanine. Five calibrants in a concentration range of 0–3.6 ng/μL (ca. 0–18 ng injected on-column) were analysed by

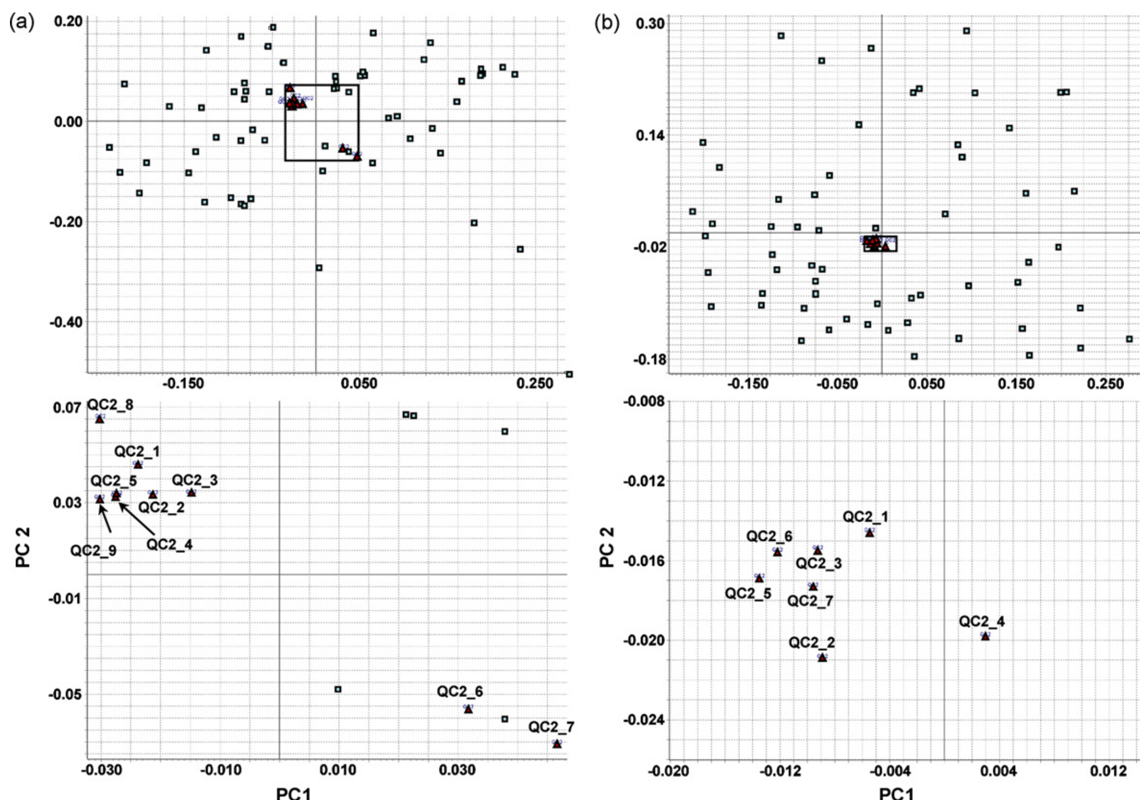


Fig. 3. Principal component analysis (mean centred and unit variance scaling) of 67 urine samples forms a lifestyle study. UPLC–TOFMS analysis was carried out in (a) positive and (b) negative ESI acquisition modes (analytical conditions are described in Section 2). Bottom figures represent an expanded view of the rectangle to zoom in the region of the QC2 samples (run number for each QC2 sample was highlighted).

UPLC–TOFMS in duplicate over a 2-day period. Two calibration curves for each analyte were realised by plotting either the peak area against the injected amount or the area ratio (analyte vs. d_5-phenylalanine) against the injected amount. A good linearity was observed (peak area against injected amount) over the concentration range measured for all analytes with respective intercept and slope values of +106.75 and 9.55 for caffeine ($r^2 = 0.9975$), of −53.59 and 35.96 for nortriptyline ($r^2 = 0.9800$), of +0.03 and 0.14 for leucine ($r^2 = 0.9609$), and of +181.64 and 4.59 for theophylline ($r^2 = 0.9901$) (positive ionisation mode). The intercept values being greater than zero reflect the presence of endogenous analytes in our urine sample (i.e., caffeine and theophylline were back calculated at respective concentration levels of 46.8 ± 1.5 μM and 176.6 ± 2.1 μM). When plotting the area ratio (analyte vs. d_5-phenylalanine) against the injected amount, the correlation coefficients were slightly lower as compared to the analyte area itself (data not shown). These results indicate that quantitative information of a particular analyte is better achieved without normalisation via the use of an internal standard (unless with a homologue isotopically labelled internal standard). It is worth mentioning that the CV obtained from the area of d_5-phenylalanine ($n = 14$, injected amount of 50 ng) was calculated at 3%, indicating a very good reproducibility of the UPLC–TOFMS system.

To complement these data, and to get a better picture of the "real-life" situation, we have analysed a urine sample spiked with the same mixture of compounds at three concentration levels 5, 10, and 15 ng/μL (ca. 25, 50, and 75 ng injected on-column) over a 7-day period. Fig. 4 depicts the area of caffeine analysed in duplicate for each day. Beside a small variation observed for the medium concentration level at day 3, a clear difference was obtained for the three different spiking levels, in a reproducible manner. Overall,

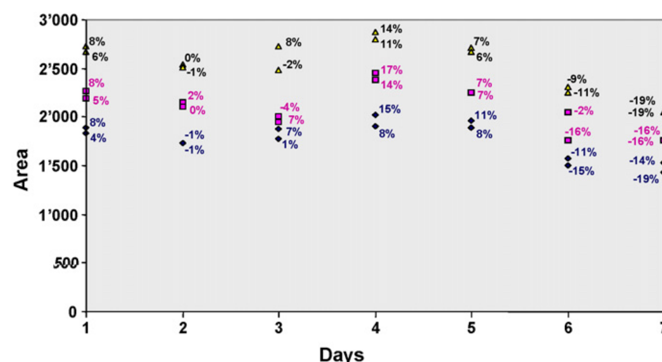


Fig. 4. The peak area of caffeine, in a spiked urine sample, was plotted over a 7-day period (injected in duplicate every day). For each data-point, the CV value was calculated from the between-day mean value of distinct concentration level (♦: 25 ng/μL, ■: 50 ng/μL, ▲: 75 ng/μL).

the CV values calculated from the analyte area of each data-point against the mean area of the corresponding spiking level were lower than 19%.

4. Conclusion

Method validation was assessed here for metabolic profiling analysis of urine samples using UPLC–TOFMS. From the chromatographic side, a longer run time of 31 min revealed greater reproducibility, and the higher number of variables was identified as compared to shortened run times (10 and 26 min). In addition, we have used two quality control urine samples originated from a distinct sample spiked with known metabolites (QC1) and from

a pooled urine sample (QC2) prepared by taking an aliquot from all urines that were collected for a clinical trial. These QC samples enabled assessment of the good quality of the data generated during the whole analytical workflow. Retention time drift, mass precision and fluctuation of the ion responses were checked over time from the QC1 runs, whereas a principal component analysis obtained from the QC2 runs, along with the unknown urine samples data, was used to ensure a tight grouping of all QC2 runs. As the most critical parameter identified was the ion response fluctuation, we proposed an acceptance criterion of ±25% (CV calculated from the mean value of spiked compounds) to accept or reject QC1 runs (urine spiked sample) to ensure a quality standard for metabolomics applications using UPLC–TOFMS.

Acknowledgements

The authors thank Dr. Ivan Montoliu Roura at the Nestlé Research Center for the critical reading of the manuscript and for the valuable comments.

References

[1] J. van der Greef, T. Hankemeier, R.N. McBurney, Pharmacogenomics 7 (2006) 1087.
[2] S. Rochfort, J. Nat. Prod. 68 (2005) 1813.
[3] I. Nobeli, H. Ponstingl, E.B. Krissinel, J.M. Thornton, J. Mol. Biol. 334 (2003) 697.
[4] W.B. Dunn, D.I. Ellis, Trends Anal. Chem. 24 (2005) 285.
[5] N. Aranibar, K.H. Ott, V. Roongta, L. Mueller, Anal. Biochem. 355 (2006) 62.
[6] H.C. Keun, T.M. Ebbels, H. Antti, M.E. Bollard, O. Beckonert, G. Schlotterbeck, H. Senn, U. Niederhauser, E. Holmes, J.C. Lindon, J.K. Nicholson, Chem. Res. Toxicol. 15 (2002) 1380.
[7] R. Stoyanova, A.W. Nicholls, J.K. Nicholson, J.C. Lindon, T.R. Brown, J. Magn. Reson. 170 (2004) 329.
[8] D.S. Wishart, D. Tzur, C. Knox, R. Eisner, A.C. Guo, N. Young, D. Cheng, K. Jewell, D. Arndt, S. Sawhney, C. Fung, L. Nikolai, M. Lewis, M.A. Coutouly, I. Forsythe, P. Tang, S. Shrivastava, K. Jeroncic, P. Stothard, G. Amegbey, D. Block, D.D. Hau, J. Wagner, J. Miniaci, M. Clements, M. Gebremedhin, N. Guo, Y. Zhang, G.E. Duggan, G.D. Macinnis, A.M. Weljie, R. Dowlatabadi, F. Bamforth, D. Clive, R. Greiner, L. Li, T. Marrie, B.D. Sykes, H.J. Vogel, L. Querengesser, Nucleic Acids Res. 35 (2007) D521.
[9] J.C. Lindon, H.C. Keun, T.M. Ebbels, J.M. Pearce, E. Holmes, J.K. Nicholson, Pharmacogenomics 6 (2005) 691.
[10] O. Fiehn, B. Kristal, B. Van Ommen, L.W. Sumner, S.A. Sansone, C. Taylor, N. Hardy, R. Kaddurah-Daouk, OMICS 10 (2006) 158.
[11] T. Ebbels, H. Keun, O. Beckonert, H. Antti, M. Bollard, E. Holmes, J. Lindon, J. Nicholson, Anal. Chim. Acta 490 (2003) 109.
[12] E. Holmes, J.K. Nicholson, G. Tranter, Chem. Res. Toxicol. 14 (2001) 182.
[13] N.J. Serkova, C.U. Niemann, Expert Rev. Mol. Diagn. 6 (2006) 717.
[14] M.E. Bollard, E.G. Stanley, J.C. Lindon, J.K. Nicholson, E. Holmes, NMR Biomed. 18 (2005) 143.
[15] J.C. Lindon, J.K. Nicholson, E. Holmes, H.C. Keun, A. Craig, J.T. Pearce, S.J. Bruce, N. Hardy, S.A. Sansone, H. Antti, P. Jonsson, C. Daykin, M. Navarange, R.D. Beger, E.R. Verheij, A. Amberg, D. Baunsgaard, G.H. Cantor, L. Lehman-McKeeman, M. Earll, S. Wold, E. Johansson, J.N. Haselden, K. Kramer, C. Thomas, J. Lindberg, I. Schuppe-Koistinen, I.D. Wilson, M.D. Reily, D.G. Robertson, H. Senn, A. Krotzky, S. Kochhar, J. Powell, O.F. van der, R. Plumb, H. Schaefer, M. Spraul, Nat. Biotechnol. 23 (2005) 833.
[16] T.J. Waybright, Q.N. Van, G.M. Muschik, T.P. Conrads, T.D. Veenstra, H.J. Issaq, J. Liq. Chrom. Relat. Technol. 29 (2006) 2475.
[17] S. Wagner, K. Scholz, M. Donegan, L. Burton, J. Wingate, W. Volkel, Anal. Chem. 78 (2006) 1296.
[18] S. Wagner, K. Scholz, M. Sieber, M. Kellert, W. Voelkel, Anal. Chem. 79 (2007) 2918.
[19] H.G. Gika, G.A. Theodoridis, J.E. Wingate, I.D. Wilson, J. Proteome Res. 6 (2007) 3291.
[20] T. Sangster, H. Major, R. Plumb, A.J. Wilson, I.D. Wilson, Analyst 131 (2006) 1075.
[21] R.E. Williams, E.M. Lenz, J.S. Lowden, M. Rantalainen, I.D. Wilson, Mol. Biosyst. 1 (2005) 166.
[22] D.J. Crockford, E. Holmes, J.C. Lindon, R.S. Plumb, S. Zirah, S.J. Bruce, P. Rainville, C.L. Stumpf, J.K. Nicholson, Anal. Chem. 78 (2006) 363.
[23] D.J. Crockford, J.C. Lindon, O. Cloarec, R.S. Plumb, S.J. Bruce, S. Zirah, P. Rainville, C.L. Stumpf, K. Johnson, E. Holmes, J.K. Nicholson, Anal. Chem. 78 (2006) 4398.
[24] P. Jonsson, S.J. Bruce, T. Moritz, J. Trygg, M. Sjostrom, R. Plumb, J. Granger, E. Maibaum, J.K. Nicholson, E. Holmes, H. Antti, Analyst 130 (2005) 701.
[25] R.S. Plumb, J.H. Granger, C.L. Stumpf, K.A. Johnson, B.W. Smith, S. Gaulitz, I.D. Wilson, J. Castro-Perez, Analyst 130 (2005) 844.
[26] A. Nordström, G. O'Maille, C. Qin, G. Siuzdak, Anal. Chem. 78 (2006) 3289.
[27] T.P. Sangster, J.E. Wingate, L. Burton, F. Teichert, I.D. Wilson, Rapid Commun. Mass Spectrom. 21 (2007) 2965.
[28] U.S. Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Veterinary Medicine (CVM), May 2001 BP. (2001) http://www.fda.gov/CDER/GUIDANCE/4252fnl.pdf.