Available online at www.sciencedirect.com
ScienceDirect
Journal of Chromatography B, 853 (2007) 20–30
JOURNAL OF CHROMATOGRAPHY B
www.elsevier.com/locate/chromb

Sample preparation and bioinformatics in MALDI profiling of urinary proteins

Panagiotis Zerefos [a], Julien Prados [b], Sophia Kossida [a], Alexandros Kalousis [b], Antonia Vlahou [a,*]

[a] *Foundation for Biomedical Research of the Academy of Athens, Athens, Greece*
[b] *University of Geneva, Geneva, Switzerland*

Received 3 November 2006; accepted 23 February 2007
Available online 19 April 2007

Abstract

One of the challenges of current proteomics research is identifying healthy and diseased mass spectrometric (MS) patterns within biological fluids. As a result, sample preparation methodologies, as well as the mathematical tools utilized for MS data analysis become pivotal. This study involves a thorough evaluation of the reproducibility and protein resolution that various urinary protein preparation methodologies provide in MALDI MS analysis. Several precipitation approaches, ultrafiltration, as well as direct dilution of urine in MALDI MS compatible buffers were applied in combination to a thorough bioinformatics analysis of the generated MS data. Our results indicate that ultrafiltration, as well as direct dilution of urine in TFA, can provide information rich and reproducible spectra for mass ranges up to 20 kDa. The importance of the presence of peak reproducibility filters when generating disease classification models is suggested.

Keywords: Urine; Proteomics; MALDI; Bioinformatics; Biomarker research

1. Introduction

The protein components of urine are useful indicators of urogenital function and human health in general. Urine samples are easily attainable, making them ideal substrates for biomarker research. Analysis of the urine proteome however, has been hindered by the great variability of the urine specimens, as well as the presence of various proteins in low abundance or modified forms. Methodologies applied for the establishment of the urinary proteome as well as the detection and analysis of disease-associated urinary peptides and proteins include two-dimensional electrophoresis [1–4], liquid chromatography coupled with mass spectrometry (LC–MS) [5,6] capillary electrophoresis in combination to mass spectrometry (CE–MS) [7] and surface enhanced laser desorption ionization MS (SELDI-TOF-MS) [8–12]. Each of these methodologies has distinct strengths and limitations and provides valuable information concerning the characterization of the urine proteome in health and disease.

The SELDI protein chip system is based on the application of protein samples on chip surfaces of distinct chromatographic properties and the subsequent analysis of the captured proteins by time of flight mass spectrometry (TOF) (reviewed in [13–15]). This methodology targets the detection of native peptides and proteins in contrast to classical LC–MS approaches that involve the analysis of proteolytic products thereof. As a result, the SELDI profiling approach does not readily provide protein identification. Nevertheless, it has distinctly high throughput capabilities. More importantly, the comparison of collected mass spectra from different sample categories (i.e. health and disease) by the use of mathematical algorithms has indicated the presence of discriminatory profiles that can classify disease with high accuracy rates [8–17]. Although highly encouraging, these initial studies pointed out various technical issues that have to be improved before any potential clinical application to this methodology could be administrated [14,16,17]. Masking of the low – by the high – abundance proteins, presence of chip to chip variability, sub-optimal mathematical tools for data processing (denoising, peak picking, intensity normalization,

* Corresponding author at: Division of Biotechnology, Foundation for Biomedical Research of the Academy of Athens, Soranou Efesiou 4, 11527 Athens, Greece. Tel.: +30 210 65 97 506; fax: +30 210 65 97 545.
E-mail address: vlahoua@bioacademy.gr (A. Vlahou).

1570-0232/$ – see front matter © 2007 Published by Elsevier B.V.
doi:10.1016/j.jchromb.2007.02.063

etc.), and detector and laser decay are considered among the major problems associated with this approach.

The study presented herein aims to unravel optimal sample preparation methodologies for the non-chip based analysis of urinary protein profiles by matrix assisted laser desorption ionization (MALDI) TOF MS. Different urine dilution, protein precipitation and ultrafiltration methods are being employed in parallel to a thorough bioinformatics analysis of the generated MS data in order to estimate the reproducibility and applicability of each approach. Our report, asides from evaluating commonly used protocols for urinary analysis by MS, also provides valuable information regarding the mathematical processing and filtering required for MS data when applied for the identification of discriminatory profiles.

2. Experimental

2.1. Materials

Amicon Ultrafiltration devices were provided by Millipore Corporation (Billerica, MA, USA), matrices for MALDI-MS: α-cyano-4-hydroxy-cinnamic (ACCA), sinapinic (SA) and 2,5-dihydroxybenzoic (DHB) acid were supplied by Sigma Corporation (St. Louis, MO, USA). Acetonitrile, ethanol, isopropanol, methanol, acetone (HPLC grade), trichloroacetic acid (TCA) and trifluoroacetic (TFA) acid (analytical purity grade) were provided by Sigma Corporation. Octyl-β-glucoside (OPG) detergent (molecular biology grade) was from Pierce Corporation (Rockford, IL, USA). NanoQ HPLC grade water, used throughout the whole experimental process as applicable, was supplied by Sarstedt Corporation (Nümbrecht, Germany).

2.2. Urine sample

A pool of random catch urine samples collected from four healthy individuals, members of the laboratory of Biotechnology (Foundation for Biomedical Research of the Academy of Athens), was utilized as standard for protocol development. All volunteers had signed an informed consent prior to the sample donation. Following their collection, urine samples were maintained at 4 °C for up to 3 h, pooled and centrifuged at 4500 RCF for 20 min for the removal of cellular components. Supernatant was aliquoted and stored at −80 °C until further use.

2.3. Urine sample preparation for MALDI MS

2.3.1. Protein precipitation

Trichloroacetic acid (TCA) precipitation was performed by adding TCA to a final concentration of 10% (w/v) to 1 mL urine followed by overnight incubation at −20 °C. The protein pellet was isolated by centrifugation at16,000 RCF/4 °C, followed by two washes with 2 mL ice-cold acetone each. Organic solvent precipitation (acetonitrile, ethanol, acetone, isopropanol, and methanol) was performed with the addition of three volumes of the respective solvent to one volume of urine, which was then followed by overnight incubation at −20 °C. The protein pellet was collected by centrifugation at16,000 RCF/4 °C, washed twice with the respective ice-cold solvent and then left to dry at ambient temperature. Pellet resuspension was performed with 200 μL of one of the following buffers: 6 M urea, 1 M guanidine, 0.1%, 0.5% or 1% TFA (v/v), 0.1% or 1% OPG (w/v), 6 M urea–0.1% OPG (w/v) solution and 0.1% TFA (v/v)–0.1% OPG (w/v). The pellets were incubated with the above buffers as applicable, at ambient temperature for 15 min, sonicated for another 15 min in ultrasound bath and diluted 10-fold with 0.1% TFA prior to their application on the MALDI plate.

2.3.2. Alternative methods for sample preparation: direct dilution – ultrafiltration

Two different dilutions as well as ultrafiltration methodologies were evaluated, as described below:

(1) Urine was diluted 10 times with 0.1% TFA (v/v) (denoted as SM).
(2) Urine was pretreated with urea at 6 M final concentration for 15 min prior to its dilution with 0.1% TFA (v/v) (SMU).
(3) Ultrafiltration was performed through Amicon filters with molecular weight cut off 5 kDa as follows: 1 mL urine was diluted to 4 mL with 0.1% TFA (v/v) and condensed via centrifugation at 3500 RCF/4 °C for 30 min. Fifty microlitre of retentate was collected and diluted 10 times with 0.1% TFA (v/v) (UF5 kDa).
(4) Ultrafiltration was conducted by the use of 10 kDa filters; in this case, the sample was pre-incubated with 6 M urea for 15 min at ambient temperature prior to its dilution with TFA and further processing as in the case of UF5 kDa (UF10 kDa).

Each of the sample preparation methodologies was conducted 10 times on different aliquots of the standard urine and each of these 10 replicates was applied five times on a MALDI target plate. One micolitre of sample plus 1 μL of matrix solution (0.5% (w/v) ACCA, saturated SA and saturated DHB in 50% acetonitrile (v/v)–0.1% TFA (v/v), as applicable) were mixed on the spot.

In parallel blank samples were processed. This involved:

1) Application of 1 μL 0.1% TFA (v/v) plus 1 μL matrix on the spot (blank to SM and SMU).
2) Filters of 5 and 10 kDa cut-off processed exactly as in the case of urine analysis but without the urine (blanks to UF5 kDa and UF10 kDa).

2.4. MS data collection

The Ultraflex I MALDI-TOF-TOF-MS (Bruker Daltonics, Bremen, Germany) instrument was used for MS data acquisition. Two acquisition modes were employed: from 1,000 to 15,000 *m/z* and 4,000 to 25,000 *m/z*. In all cases, spectra accumulation was performed by adding 1000 laser shots in positive linear mode (100 laser shots at50 Hz repetition rate) and 3 kV accelerating voltage. Calibration of the spectra was performed externally by employing ClinProt Calibration standard and Protein Calibration Standard I from Bruker Daltonics. Deposition

of the samples on the MALDI target and mass spectra acquisition were performed randomly. Following their collection, the raw mass to charge data was extracted as ASCII files for further feature extraction by the use of the mathematical algorithms described below.

2.5. Pre-processing of MS data

A novel peak detection algorithm has been used for signal processing. The efficiency of the algorithm in MS data processing has been recently demonstrated elsewhere [18]. In brief, the algorithm defines peaks as all points p of a spectrum whose left and right valleys are deeper than the noise level at p. More specifically, the algorithm establishes its corresponding left and right valleys for each point p of a spectrum. These are defined as the minimum intensity point on the left and right of p, respectively, so that there is no point with intensity higher than the intensity of p between the two valleys. The minimum intensity difference between the point and its left and right valleys corresponds to the intensity of the peak located at p. It should be noted that the algorithm sets no thresholds on the width of a peak which allows the detection of peaks of varying width within a spectrum. For peak detection, smoothing and estimation of valley depth and noise levels were performed. Specifically, smoothing is applied on the original raw spectrum using a moving average window with size of 30 data points. Estimation of the valley depth involves computing the smallest of the two valley depths of each point of the spectrum. The estimation of the noise level is obtained by subtracting the smoothed signal from the original signal and then computing a moving local standard deviation within a neighbourhood of 1000 data points. It should be noted that peak detection, as described above, is not baseline sensitive, since peaks are detected relative to valleys within their neighbourhood which have similar baseline levels. This strategy of peak detection is preferable for the analysis of complicated and "noisy" signals, where noise and baseline levels are difficult to estimate precisely [18].

For the estimation of peak intensities, baseline removal and total ion current (TIC) normalization are further applied. For baseline removal, the baseline with a closing filter over 1/20th of the total number of data points is computed and subtracted from the input (a closing filter is a filter which initially finds the moving local minima and then identifies the local maxima of the local minima for each point of the input signal). TIC normalization is performed according to total ion current, meaning that the signal is rescaled by dividing it with the sum of its intensities.

An example of the application of these pre-processing methods on a spectrum is given on Fig. 1. Peaks are located at positions where the blue bars plot (valley depth) is above the green line (noise level). The ratio between the two corresponds to the signal/noise ratio (S/N) of the peaks. It should be noted that the valley depths do not faithfully reflect the intensity of the peaks in the cases of overlapping peaks. Hence, the difference in intensity between the baseline (dashed blue line) and the signal was preferred to the valley depth as an estimate of the peak intensities.

2.6. Comparison of MS data

For the comparison of the peaks included in different spectra, peak alignment was performed with an agglomerative hierarchical clustering algorithm [19]. In brief, the algorithm initially creates as many clusters as detected peaks in all spectra and successively merges the closest clusters if: (1) all the peaks of

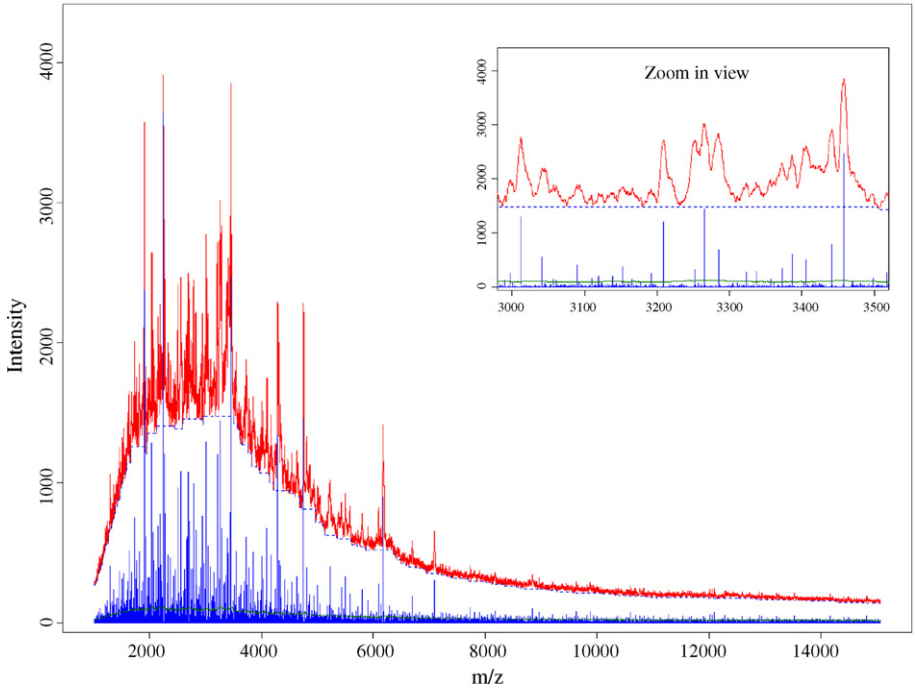

Fig. 1. Illustration of peak detection. Total mass spectrum view and zoom in the 3000–3500 m/z area showing the smoothed spectrum (in red), the noise level (in green) and the baseline (dashed blue line). The blue histogram corresponds to the valley depth of each point.

the new cluster belong to different spectra, and (2) the maximal mass distance between two peaks of the new cluster is less than 0.2% of their mean. This tolerance corresponds to a measurement error of 0.1%. In detail, if min and max are the minimum and maximum masses of a new potential cluster, and M is their mean ($M = (\min + \max)/2$), then the cluster is retained only if ($\max - M < M \times 0.1\%$) and ($M - \min < M \times 0.1\%$) which is equivalent to considering the interval ($M \pm M \times 0.1\%$) of size $M \times 0.2\%$. The manufacturers of the MALDI-TOF instrument (Bruker Daltonics) recommend this specific mass error.

In order to establish the similarity of any two methods on the basis of the peak sets that they generate, a measure of similarity or distance between sets was calculated using the Tanimoto set distance. Specifically, if X and Y are two sets of peaks, the Tanimoto distance is given by:

$$D(X, Y) = 1 - \frac{|X| + |Y| - 2|X \cap Y|}{|X| + |Y| - |X \cap Y|}$$

Hierarchical clustering with complete linkage was then applied in order to determine if and how the different methods cluster together.

2.7. Statistical analysis

Statistical analysis was performed for the signal intensities of the peaks detected. The coefficient of variation (CV) of the intensities was calculated and correlated to the mean intensity. This correlation was estimated based on Pearson's product moment correlation coefficient which follows a t-distribution. For $p < 0.05$, the null hypothesis "correlation coefficient is zero" is rejected.

3. Results and discussion

3.1. Sample preparation and MALDI MS

The central idea of this study is the clarification of sample preparation and bioinformatics issues involved in urinary TOF MS profiling. In the case of serum protein profiling by MS, there is a consented effort to resolve some of these issues and establish a reproducible and portable profiling assay for the detection of prostate cancer [17]. Compared to serum, urinary profiling by MS is still at its infancy. Nevertheless, several reports support the potential of this approach as a diagnostic tool for urogenital disease, emphasizing the need for the establishment of standardized procedures for urine sample storage, protein processing as well as MS data analysis [8–12].

The initial studies on MS urinary profiling relied on the use of the chromatographic chip surfaces employed by the SELDI technology (Ciphergen Biosystems). Application of the chips allows for sample enrichment of specific proteins according to their chemical properties and the type of surface in use. However, chip surfaces are a significant source of variability. Slight variations in the chemical properties between chip batches result in significant differences in the quality of the generated MS data (reviewed in [13–15]).

With the study presented herein, a thorough evaluation of the non-chromatographic standard MALDI chip for urinary profiling is performed. In the case of MALDI profiling, protein samples are simply dried on the surface prior to their analysis by MS.

We used the standard MALDI plate surface in combination with various classical urinary protein enrichment methodologies. At this initial screening phase, each method was repeated at least three times and was evaluated based on the quality of the MS output (i.e. number, resolution, intensities and reproducibility of resolved peaks) generated through the Flex Analysis 2.2 software (Bruker, Daltonics). Among others, protein precipitation approaches by the use of TCA or organic solvents (acetone, ethanol, acetonitrile, isopropanol and methanol) were employed. Prior to MALDI MS analysis, the protein pellets were solubilized with several chemicals including urea, guanidine, trifluoroacetic acid, OPG detergent or combinations thereof, as described in the experimental section. The respective supernatant solutions (containing the smaller and soluble biomolecules), with the exception of the TCA precipitation supernatant, were also subjected to MALDI MS analysis. The quality of the generated MS spectra (Fig. 2, I–II) following the use of the above precipitation and solubilization means was either considered quite poor, or the method presented serious reproducibility problems. This result cannot be attributed to protein loss since the same precipitated samples could be very well and reproducibly resolved by the use of 1 and 2D electrophoretic approaches (data not shown). Without ruling out the possibility that a potentially different sample treatment might have generated superior MS spectra, we concluded from this analysis, that traces of the chemicals utilized or retained (for example plasticizers) during precipitation or solubilization have a significant suppression effect on the ionization process.

To our surprise, direct dilution of the urine sample in 0.1% TFA or urea with TFA provided better MS spectra for 1,000–15,000 m/z, (Fig. 2, III–IV SM and SMU). The selection of these chemicals was based on their properties. Urea is considered to act as a chaotropic agent facilitating protein solubilisation and denaturation. Even strong protein aggregates or complexes can be dissolved with the use of urea solutions higher than 1 M. On the other hand, TFA, like many fluorinated solvents, is a well known solubilising agent. TFA acts as an acid and an ion pairing reagent, while its volatility permits its use (at low concentrations) in MALDI-MS.

Similarly, urine samples subjected to ultrafiltration by the use of 5 or 10 kDa cut-off filters generated relatively rich protein spectra for the same mass range (Fig. 2, III–V UF5 kDa, UF10 kDa). These specific cut-offs were chosen in order to minimize inhibition effects in spectrum acquisition from salts and small molecules. It is also evident from the generated spectra that although an elimination of the biomolecules with MW below 5 or 10 kDa, respectively, should have been theoretically performed, this does not occur; low molecular weight peptides and proteins are retained and get ionized very efficiently. The 10 kDa filters were used in combination to urea. This was done in order to facilitate contaminant removal, suppress biomolecule clustering and facilitate ionization of larger proteins and polypeptides.

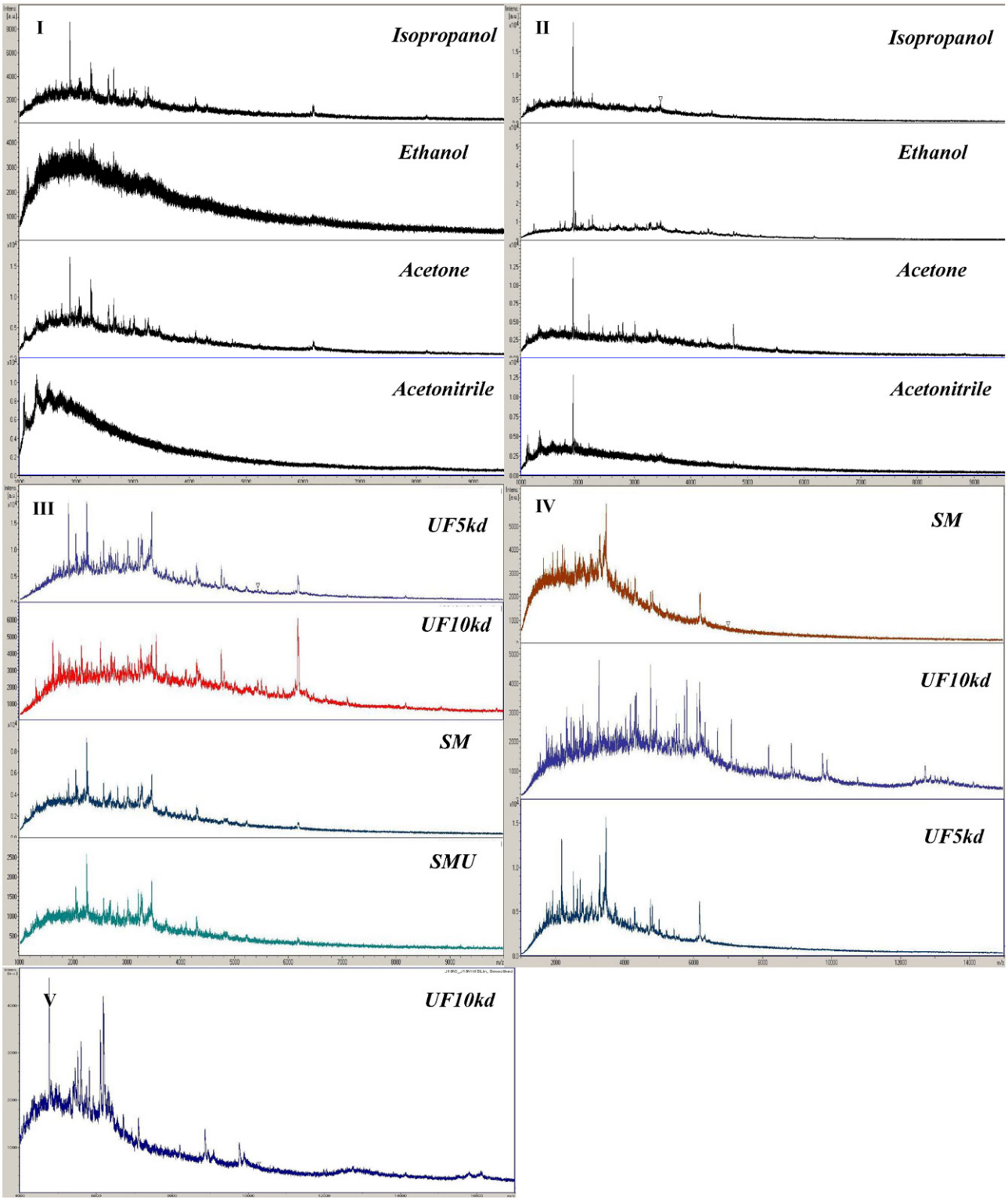

Fig. 2. Representative spectra collected by the use of the precipitation, ultrafiltration and direct dilution methods, and visualized by the use of the MALDI Flex Analysis 2.2 software. (I–II): Spectra collected following application of organic solvent precipitation means (as indicated) from: I: the pellets dissolved in TFA and II: the respective supernatants. In both I and II, the ACCA matrix was employed. These methods provided either poor MS spectra (I ethanol, isopropanol) or presented serious reproducibility problems. (III–V): Spectra collected following the application of the ultrafiltration (UF5 kDa, UF10 kDa) and direct dilution (SM, SMU) methods in combination to: (III): ACCA; (IV): DHB and (V): SA. *x*-axis: *m/z* (1,000–10,000 for I–IV, 4,000–20,000 for V), *y*-axis: signal intensity.

Various others chaotropic solutions were also evaluated for the same purposes such as guanidine, acetonitrile, acetonitrile with isopropanol and OPG detergent in MS compatible concentrations. From this analysis urea was found to generate the most informative and less noisy (from interfering substances such as plasticizers) spectra (data not shown).

Collectively, based on this initial screening, we opted to proceed with a thorough evaluation of the reproducibility and peak information provided by the two different dilution methods of the starting urine sample (in 0.1% TFA (v/v)–SM and urea–SMU) as well as the two ultrafiltration methods (UF5 kDa, UF10 kDa). These methods were utilized in combination with different widely utilized matrices, namely ACCA and DHB. Both of these matrices are applicable for peptide and low MW protein ionization and have distinct energy transfer capabilities. The combination of urea with DHB was avoided since the former strongly inhibits crystal formation and suppresses ionization triggered by the latter. Sinapinic acid was utilized only in combination to the UF10 kDa, which targets the ionization and detection of higher MW proteins.

Two different acquisition methods were applied for *m/z* regions from 1,000 to 15,000, and 4,000 to 25,000, respectively. These two acquisition methods were selected in order to maximize the efficiency of the matrices employed in protein desorption. Specifically, the 1000–15000 *m/z* acquisition mode was used, in cases where ACCA and DHB were employed; SA is known to facilitate the ionization of larger MW proteins and thereby it was used in combination to the 4000–25000 *m/z* acquisition mode. In addition, the lower limit of 1000 *m/z* was set to eliminate peaks that may belong to matrix and other non-protein interfering molecules. The upper limit of 25,000 *m/z* in the detection mode was set primarily based on the sensitivity and the resolving power of the MALDI-TOF instruments. These parameters are optimal for low molecular weight polypeptides and decline with molecular weight increase. This is attributed to the fact that the ionization efficiency of biopolymers (proteins, DNA and carbohydrates) is reduced with size increment. The general rule is that the larger the molecules, the greater the amount of energy that is required in order to reach the gaseous phase. Along the same lines, the vast majority of disease discriminatory peaks and peak profiles detected by SELDI/MALDI profiling and described in the literature are in the <15000 *m/z* range (reviewed in [14]). In any case, we should note that by the use of different sample preparation or separation methodologies, mass spectrometric profiling of high molecular weight proteins might be feasible. However, in our opinion, it is of greater importance to optimize and standardize the methodology at the low (<20,000) *m/z* range since: (a) as mentioned, in this low mass range the instrument exhibits greatest sensitivity in peptide/protein detection and (b) this is a mass range that gel-based systems do not efficiently and easily resolve.

3.2. Reproducibility in peak detection

Each sample preparation method was repeated 10 times in different aliquots of the starting material and each of these 10 replicates was applied in quintuplicate on the MALDI target. The former measurements serve to estimate the between-run or method precision whereas the latter the within-run precision. Blank samples for each method were similarly processed to offer an estimate of the non-specific background interference affecting the assay.

3.2.1. Within-run precision

Peak detection and alignment were performed for each of the five replicate sample applications of a given method, in order to identify the distinct peaks among the five replicates. The number of peaks detected in exactly i replicates, where $i := 1 \ldots 5$, for different S/N ratios (from 0.6 to 6 at increments of 0.2) was then calculated and expressed as the relative percentage of the total number (i.e. from all five replicates) of detected peaks.

The results for the ACCA/UF5 kDa and ACCA/UF10 kDa preparation methods for one random set of their quintuplicate applications and their respective blanks are shown in Fig. 3. Of the two methods, ACCA/UF5 kDa is by far the more reproducible since for a wide range of S/N threshold values (from 3 up to 6), more than 40% of the detected peaks are found systematically in all five replications, and roughly 50–60% of the detected peaks are found in 4/5 replicates. These percentages refer to cumulative data; i.e. in a total of 172 peaks, 72 (41.9%) peaks are detected in 5/5 replicates of ACCA/UF5 kDa. For the same S/N values, the corresponding reproducibility values for ACCA/UF10 kDa are less than 10 and 20%, respectively. This result is representative of the reproducibilities received by any of the 10 quintuplicate sets of the two methods (Table 1 of the appendix). The level of reproducibility by ACCA/UF5 kDa is matched only by the SA/UF10 kDa method for S/N > 3, and followed by the ACCA/SM method with about 35% of peaks detected in 5/5 replicates for S/N > 3. In general, employment of ACCA provided superior reproducibility compared to the DHB for almost all tested S/N values.

Peaks are also detected within the blank samples (Fig. 3). These peaks are usually attributed to the matrix and other interference substances (for example plasticizers) Nevertheless, it is important to note that in contrast to the urine samples, the number and reproducibility of the 'background' peaks significantly decreases for S/N > 3.4 (Fig. 3). In this case, <2 peaks were detected in at least 3/5 blank replicates for the vast majority of methods (Table 7 in the appendix).

3.2.2. Between-run precision

In the comparison of the 10 replicate sets of each methodology, the peaks present in at least 3/5 replicates within each set were utilized. These 'robust' peaks were aligned (Fig. 4) and the number of peaks detected in exactly i runs ($i := 1 \ldots 10$) of the method, as well as the respective percentages over the cumulative number of peaks detected with each method, were calculated.

The results of this analysis are depicted in Fig. 5. In accordance to the within-run precision analysis, for ACCA/UF5 kDa approximately 50% of the detected peaks are systematically present in all 10 runs of the method for S/N threshold values between 2 and 6; this percentage increases to 60% if we consider the number of peaks present in at least 9/10 replicates.

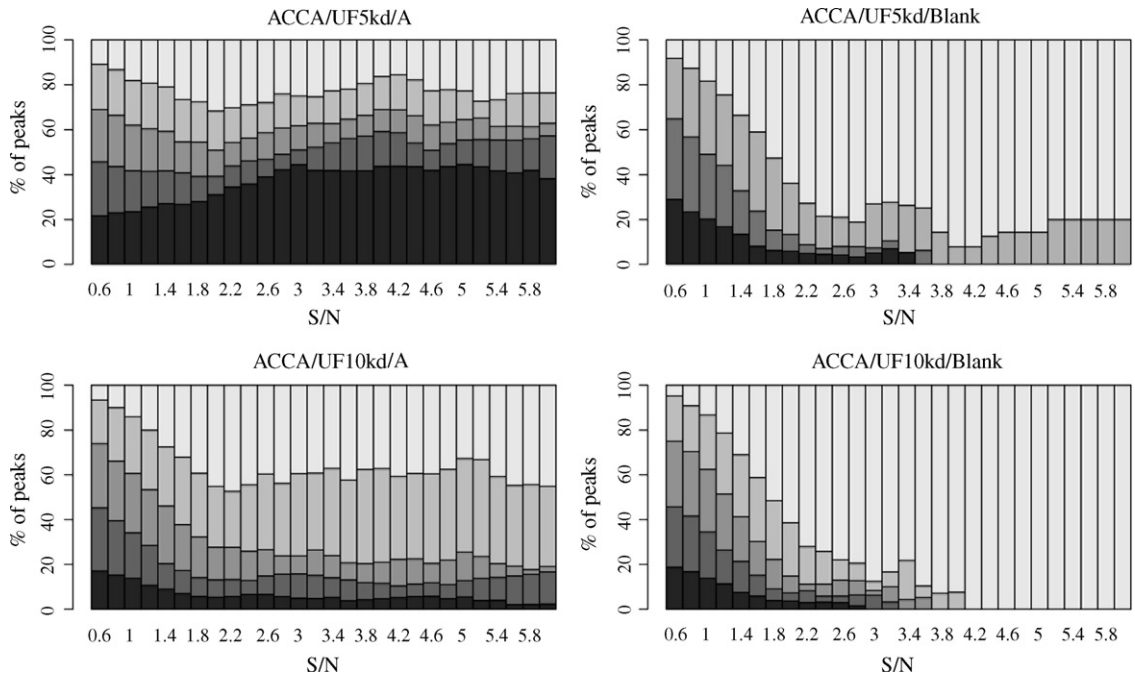

Fig. 3. Within run peak reproducibility for one of the quintuplicate sets of ACCA/UF5 kDa, ACCA/UF10 kDa and respective blanks. The percentage of peaks detected in 5/5 (darkest), 4/5, 3/5, 2/5, and 1/5 (lightest) sample applications on the target at different S/N ratios is shown. Measurements for the blank sample are also shown (background inference is negligible for S/N > 3).

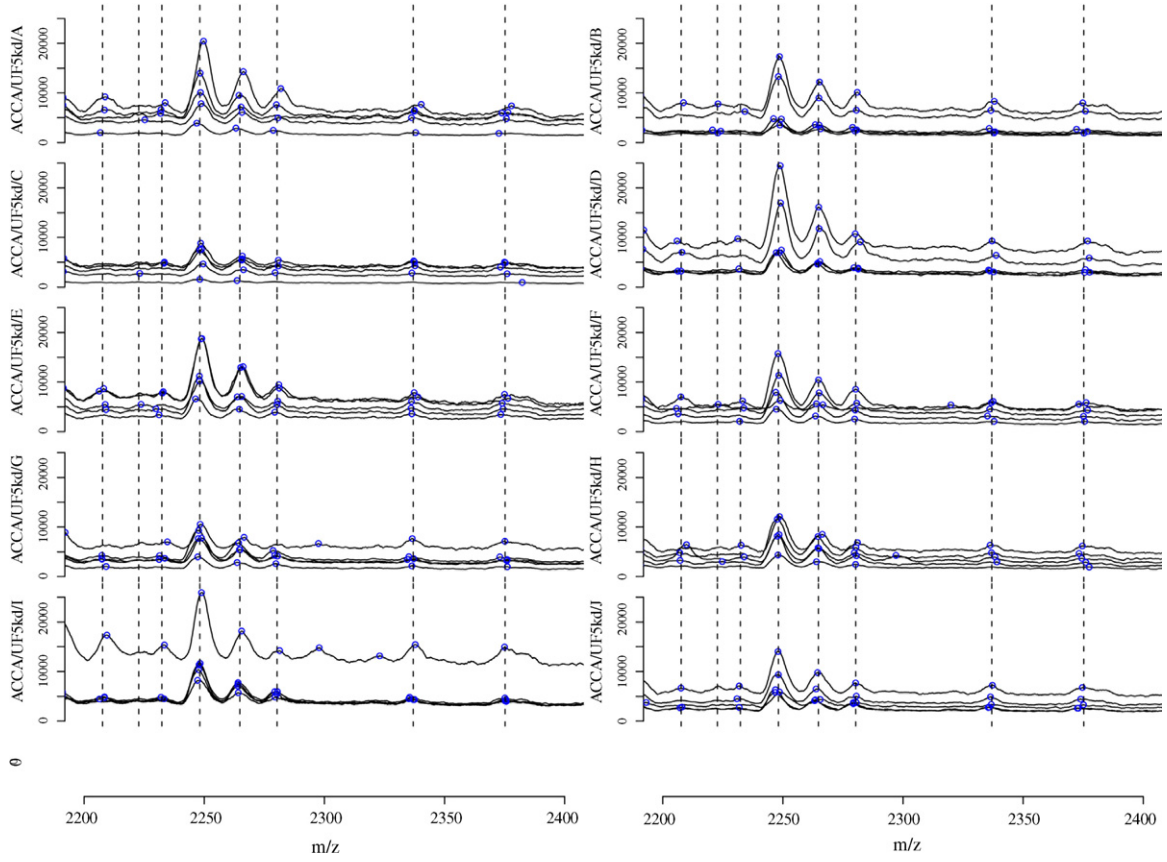

Fig. 4. ACCA/UF5 kDa spectra alignment. The 50 spectra of ACCA/UF5 kDa separated to the 10 repetitions of the method are shown. The blue points correspond to the peaks detected in each spectrum. Vertical lines depict the final set of peaks after alignment over all the 10 repetitions of the given method. Mass to charge range shown: 2200–2400.

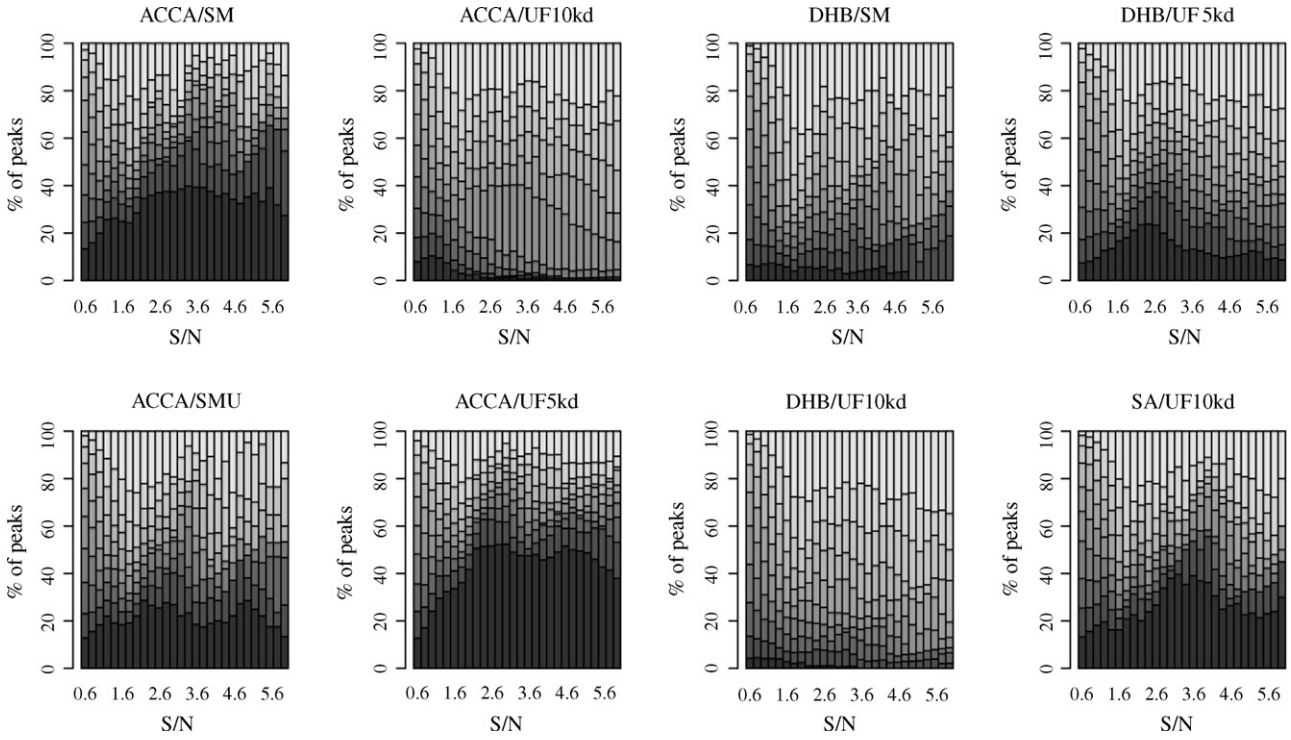

Fig. 5. Between run peak reproducibility of the different methods. Percentage of peaks detected in 10/10 (darkest) up to 1/10 (lightest) repetitions of each sample preparation method is shown.

For ACCA/SM, the respective values are 40 and 55%. Similarly for SA/UF10 kDa, 40 and 50% of peaks are detected in 10/10 and 9/10 replicate sets, respectively, for more narrow S/N ranges (3–3.6). ACCA/SMU and DHB/UF5 kDa exhibit significantly lower reproducibilities. In all cases, it is noteworthy that the between-run precision is very similar to the within-run precision. This suggests that the main source of variation is not related to the sample preparation procedure per se but rather to the ionization process and instrument performance. The latter is a significant source of variation as stated in previous reports [16,17]. It should also be noted that in our case, matrix addition was performed manually, potentially contributing to the observed variation.

3.2.3. Number of detected peaks

Besides the reproducibility in peak detection, the different methods were also compared for the number of peaks they resolved. To perform this analysis, peaks present in at least 3/5 replicates within each set and at S/N of 3.4 were considered. The latter decision was based on the results of the within-run precision analysis that indicated that at this S/N ratio background peaks are negligible.

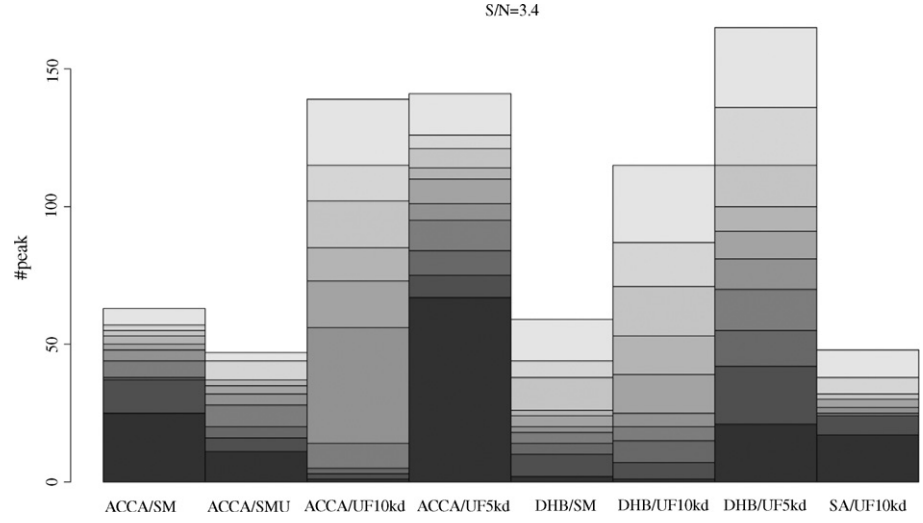

Fig. 6. Number of peaks detected in 10/10 replicates (dark grey) down to 1/10 (light gray) replicate sets of the different methods.

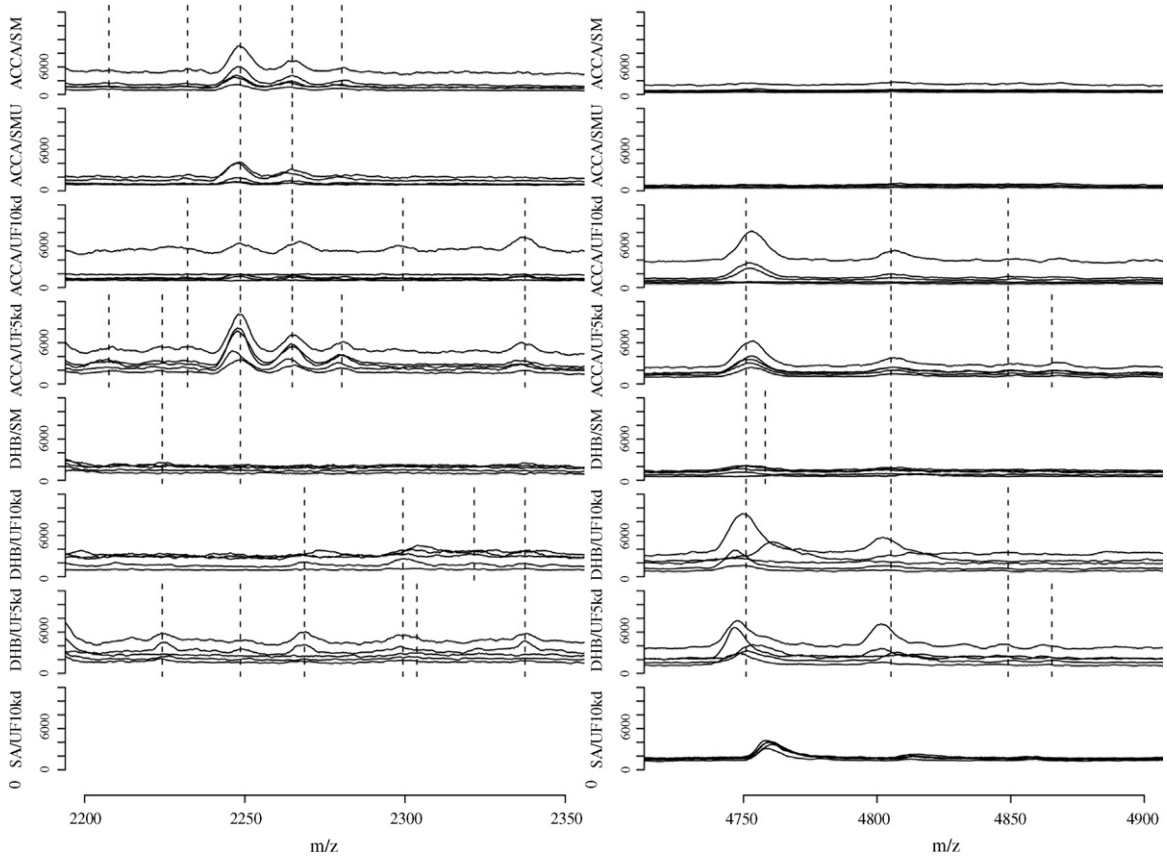

Fig. 7. Spectra alignment along three different *m/z* regions (columns) of the different methods (rows). For each method, five randomly selected spectra are shown. Vertical lines correspond to the aligned peak locations; if the lines are not present; the corresponding peak was not detected in the spectra of that method. S/N threshold value is 3.4. Mass to charge range shown: 2200–2250 and 4750–4900.

ACCA/UF5 kDa is the method which provides the highest number of detected peaks; specifically the median number of detected peaks in the 10 replicates (± standard deviation) is 107 (±8). It is also the method that exhibits the highest number of reproducible peaks with 67/141 peaks (47.5%) detected in all 10 replicate sets (Fig. 6 and Table 3 of the appendix). DHB/UF5 kDa and ACCA/UF10 kDa also show a relatively high number of detected peaks with median (±S.D.) values of 90 (±24) and 80 (±36); DHB/UF10 kDa, ACCA/SM and SA/UF10 kDa overall generate lower numbers of detected peaks with median (±S.D.) values of 50 (±25), 48 (±6), and 30 (±2), respectively, (Fig. 6 and Table 3 of the appendix).

3.3. Method complementarity

To evaluate the complementarity in peak information that the different methodologies provide, alignment of peaks detected in different methods was performed. For this type of analysis, the possibility of potential mass shifts generated by the different methodologies had to be tested. The steps followed for this analysis have been described above; i.e. an initial alignment of the peaks detected in the five replicate applications is performed. This step is followed by the elimination of the peaks that were not detected in at least 3/5 replicates. Subsequently, the robust peaks found in at least 3/5 replicates are aligned in the 10 repetitions of a method. A final alignment is then performed along the different methods. This process is shown in Fig. 7 for two different *m/z* regions and five randomly selected spectra for each method. Based on visual inspection, no serious shifts of masses are observed between the different methodologies, with the possible exception of the SA/UF10 kDa, which was therefore excluded from the subsequent clustering analysis.

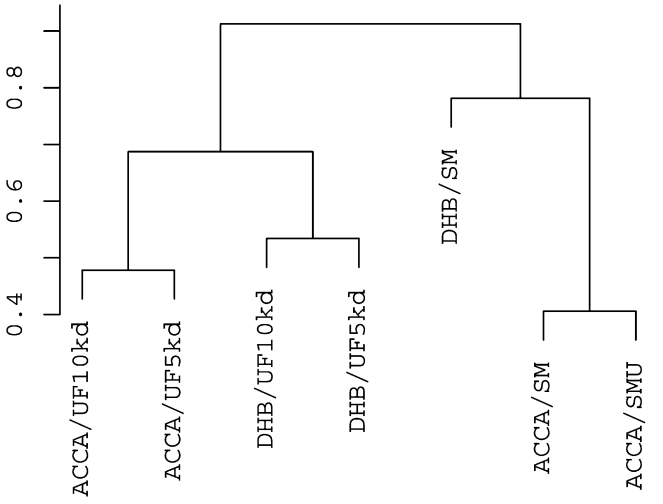

Fig. 8. Clustering of the different sample preparation methods according to the number of peaks they have in common. The distance between horizontal lines is a measure of similarity.

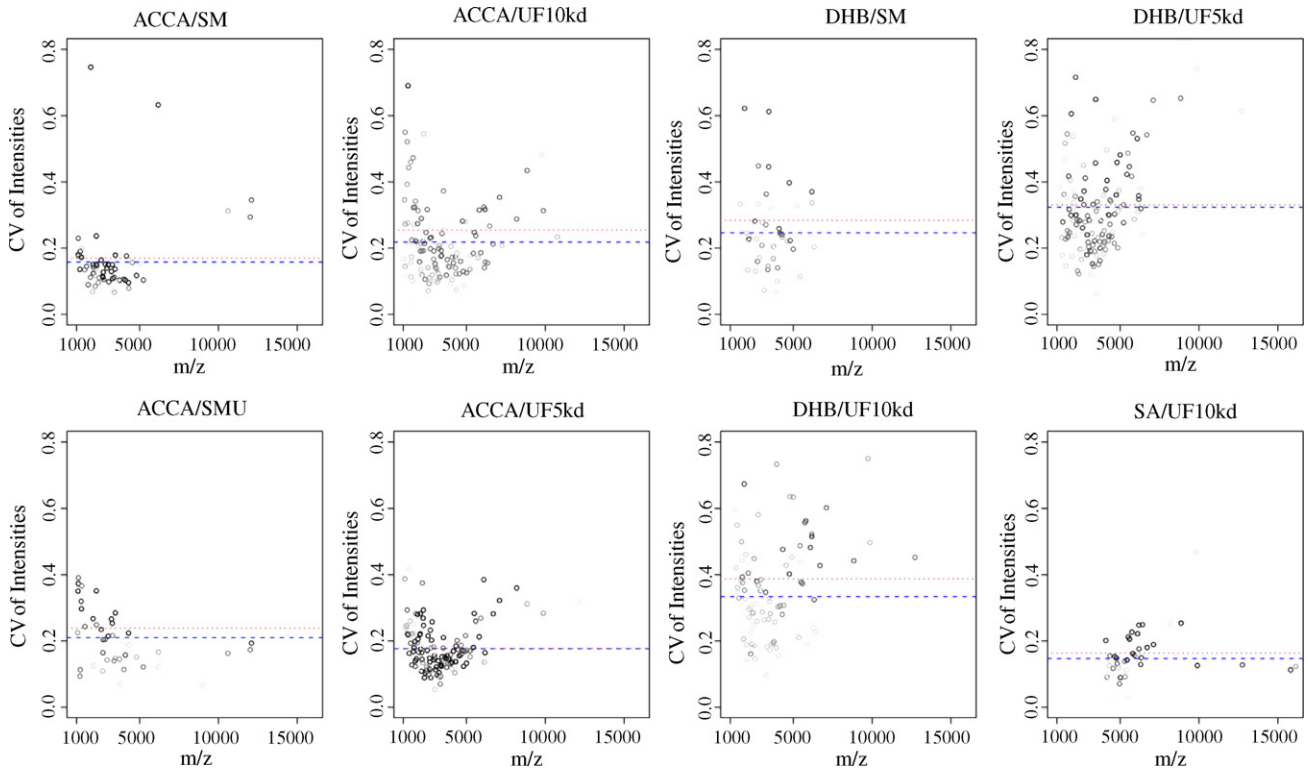

Fig. 9. Coefficient of variation of peak intensities for the different peaks detected by each method. S/N threshold was set to 3.4. Each point corresponds to a detected peak; the intensity of its colour corresponds to the number of times that the peak was detected in the 50 different spectra of a given method (the darker the colour the more spectra it was detected in). *x*-axis: *m/z*; *y*-axis: CV of peak intensities among the 50 spectra. The mean CVs among all detected peaks are shown (horizontal blue line), as well as the respective weighted mean (weighted by the number of times a peak was detected; red line).

The results of hierarchical clustering with complete linkage of the different methods, based on the peak information they provide, are shown in Fig. 8. As expected, similar methods cluster together (for example ultrafiltration methods), whereas the complementarity of different methods and matrices is observed (for example ACCA/SM methods with DHB/UF methods are highly complementary; detailed data on the number of common peaks are given in Table 4 of the appendix). Considering the fact that overall a lower number of peaks are resolved in urine samples compared to serum by MS [14–17], an analysis of the former by the use of a combination of different reproducible and complementary approaches is recommended as a means to increase protein resolution.

3.4. Reproducibility in peak intensities

So far, the reproducibility of the different methods was evaluated based on the presence/absence of peaks. Central to the profiling assays however is the reproducibility or better stated variation of peak intensities. To evaluate the different urine preparation methodologies for this factor, peaks were detected as previously described for the S/N ratio of 3.4. In addition, baseline correction and intensity normalization according to TIC were performed. The CV of peak intensities was then calculated and the results are provided in Fig. 9.

Based on this analysis, the methods with the best peak reproducibility were also found to exhibit the lowest intensity CVs. Specifically, ACCA/UF5 kDa, ACCA/SM and SA/UF10 kDa exhibit CVs of less than 20%. These values are similar and even superior to intensity CVs regularly observed in SELDI/MALDI profiling assays [16,17,20]. It can also be seen that intensity CVs do not relate to the peak *m/z* values (Fig. 9). In addition, for these three methods, no correlation was found between the intensity CV and the mean intensity value of a peak (p-value > 0.05, Table 6 of the appendix). Simply stated, the variation in the intensity of a peak is independent of the absolute value of its intensity and its *m/z*.

4. Conclusion

In the recent past, several studies have been conducted evaluating urinary protein preparation methods prior to electrophoretic or chromatographic separation. However, there is a lack of similar information when direct mass spectrometric profiling of urine is employed. The novelty of our report is that it provides a coherent and thorough analysis of the reproducibility and overall performance of different classical urinary protein preparation methodologies in MALDI TOF MS profiling. Although the latter approach does not provide protein identification, it has received much attention lately due to (a) its high throughput capabilities, (b) its ability to resolve low molecular weight native peptides and proteins and (c) the high accuracy rates in disease detection provided by the resolved proteomic or better stated ionic patterns. In our study, several novel elements are presented. Evaluation of urine protein analysis by standard MALDI chips is performed in contrast to the

by now analysis on chromatographic chip surfaces (SELDI). A thorough analysis of the performance of a variety of classical approaches for urinary protein preparation such as precipitation by various means, ultrafiltration and direct urine dilution is performed. Given the relatively large number of available replicates (50), several quality control characteristics are assessed such as the within- and between-run precision of peak detection, the intensity CV throughout the mass spectrum and its correlation to *m*/*z* and to the individual peak intensity. Our results indicate that ultrafiltration methods (ACCA/UF5 kDa and SA/UF10 kDa) as well as direct dilution of urine to MS compatible buffers such as 0.1% TFA can provide reproducible peak information. It should be noted that day-to-day variations or variations attributed to sample handling (i.e type of urine collection tube, time of sample at 4 °C, etc.) have not been addressed in this study. However, these issues have been nicely described elsewhere [8]. What is evident from our study however, is that methodologies that appear to provide relatively rich mass spectra, such as the ACCA/UF10 kDa or DHB/UF5 kDa may fail in extensive reproducibility studies. Apparently, disease classification models generated by these approaches are deemed to fail as well. Hence, based on the results presented herein, we would favor the use of stringent criteria during the peak detection and bioinformatics analysis of MS data from clinical samples. These criteria include: a relatively high number of replicates per sample (at least 4), the employment of S/N thresholds where background peaks significantly reduce (3.4 in our case) and the elimination of the non-reproducible peaks. This approach may be eliminating some low-abundance peaks; however, it is more prone to succeed in reproducibility and portability studies and consequently in generating robust disease classification models.

Acknowledgement

This work was supported financially by the Greek Ministry of Health.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jchromb.2007.02.063.

References

[1] R. Pieper, C.L. Gatlin, A.M. McGrath, A.J. Makusky, M. Mondal, M. Seonarain, E. Field, C.R. Schatz, M.A. Estock, N. Ahmed, N.G. Anderson, S. Steiner, Proteomics 4 (2004) 1159.
[2] J.X. Pang, N. Ginanni, A.R. Dongre, S.A. Hefta, G.J. Opiteck, J. Prot. Res. 1 (2002) 161.
[3] P. Zerefos, K. Vougas, P. Dimitraki, S. Kossida, A. Petrolekas, K. Stavrodimos, A. Giannopoulos, M. Fountoulakis, A. Vlahou, Proteomics 6 (2006) 4346.
[4] V. Thongboonkerd, S. Chutipongtanate, R. Kanlaya, J. Prot. Res. 5 (2006) 183.
[5] M. Soldi1, C. Sarto1, C. Valsecchi1, F. Magni, V. Proserpio1, D. Ticozzi, P. Mocarelli, Proteomics 5 (2005) 2641.
[6] M. Jurgens, A. Appel, S. Neitz, C. Menzel, H. Tammen, H.D. Zucht, Comb. Chem. High Throughput Screen. 8 (2005) 757.
[7] P. Zurbig, M.B. Renfrow, E. Schiffer, J. Novak, M. Walden, S. Wittke, I. Just, M. Pelzing, C. Neususs, D. Theodorescu, K.E. Root, M.M. Ross, H. Mischak, Electrophoresis 27 (2006) 2111.
[8] M.A. Rogers, P. Clarke, J. Noble, N.P. Munro, A. Paul, P.J. Selby, R.E. Banks, Cancer Res. 15 (2003) 6971.
[9] A. Vlahou, P.F. Schellhammer, S. Mendrinos, K. Patel, F.I. Kondylis, L. Gong, S. Nasim, G.L. Wright, Am. J. Pathol. 4 (2001) 1491.
[10] A. Vlahou, A. Giannopoulos, B.W. Gregory, T. Manousakas, F.I. Kondylis, L.L. Wilson, P.F. Schellhammer, G.L. Wright, O.J. Semmes, Clin. Chem. 8 (2004) 1438.
[11] N.P. Munro, D.A. Cairns, P. Clarke, M. Rogers, A.J. Stanley, J.H. Barrett, P. Harnden, D. Thompson, I. Eardley, R.E. Banks, M.A. Knowles, Int. J. Cancer 119 (2006) 2642.
[12] Y.F. Zhang, D.L. Wu, M. Guan, W.W. Liu, Z. Wu, Y.M. Chen, W.Z. Zhang, Y. Lu, Clin. Biochem. 37 (2004) 772.
[13] K. Chapman, Biochem. Soc. Trans. 30 (2002) 82.
[14] E.P. Diamandis, Mol. Cell. Proteomics 3 (2004) 367.
[15] A. Vlahou, M. Fountoulakis, J. Chromatogr. B 814 (2005) 11.
[16] A.J. Rai, P.M. Stemmer, Z. Zhang, B.L. Adam, W.T. Morgan, R.E. Caffrey, V.N. Podust, M. Patel, L.Y. Lim, N.V. Shipulina, D.W. Chan, O.J. Semmes, H.C. Leung, Proteomics 5 (2005) 3467.
[17] O.J. Semmes, Z. Feng, B.L. Adam, L.L. Banez, W.L. Bigbee, D. Campos, L.H. Cazares, D.W. Chan, W.E. Grizzle, E. Izbicka, J. Kagan, G. Malik, D. McLerran, J.W. Moul, A. Partin, P. Prasanna, J. Rosenzweig, L.J. Sokoll, S. Srivastava, I. Thompson, M.J. Welsh, N. White, M. Winget, Y. Yasui, Z. Zhang, L. Zhu, Clin. Chem. 51 (2005) 102.
[18] J. Prados, A. Kalousis, M. Hilario, 9th IEEE International Symposium on Computer-Based Medical Systems, (2006) 953.
[19] J. Prados, A. Kalousis, J. Sanchez, L. Allard, O. Carrette, M. Hilario, Proteomics 8 (2004) 2320.
[20] L.J. Dekker, J.C. Dalebout, I. Siccama, G. Jenster, P.A. Sillevis Smitt, T.M. Luider, Rapid Commun. Mass Spectrom. 19 (2005) 865.